



**HAL**  
open science

# Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans le diagnostic des besoins langagiers en FLE à l'aide des techniques de TAL

Rui Yan

## ► To cite this version:

Rui Yan. Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans le diagnostic des besoins langagiers en FLE à l'aide des techniques de TAL. Linguistique. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAL007 . tel-01691923v2

**HAL Id: tel-01691923**

**<https://theses.hal.science/tel-01691923v2>**

Submitted on 24 Jan 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.

## **THÈSE**

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

Spécialité : **Sciences du langage Spécialité Didactique et  
Linguistique**

Arrêté ministériel : 25 mai 2016

Présentée par

**« Rui/ YAN »**

Thèse dirigée par **Agnès TUTIN** et  
codirigée par **Cristelle CAVALLA**

préparée au sein du **Laboratoire LIDILEM**  
dans **l'École Doctorale Langues, Littératures et Sciences  
Humaines**

### **Étude des constructions verbales scientifiques dans une perspective didactique: utilisation des corpus pour le diagnostic des besoins langagiers en FLE à l'aide des techniques de TAL**

Thèse soutenue publiquement le « **26 juin 2017** »,  
devant le jury composé de :

**M. Francis GROSSMANN**

Professeur, Université Grenoble Alpes, Président

**M. Jan GOES**

Professeur, Université d'Artois, Rapporteur

**Mme. Natalie KÜBLER**

Professeur, Université Paris Diderot, Rapporteur

**Mme. Agnès TUTIN**

Professeur, Université Grenoble Alpes, Directrice

**Mme. Cristelle CAVALLA**

Maître de conférences, Université Sorbonne Paris 3, Co-directrice

**M. John OSBORNE**

Professeur, Université de Savoie Mont Blanc, Examineur



---

---

## RÉSUMÉ :

La présente étude est centrée sur les constructions verbales scientifiques. Ce lexique se situe à mi-chemin entre le lexique de la langue générale et celui de la terminologie. Il s'agit de termes utilisés à des fins d'expressions scientifiques, pour contribuer à l'élaboration du fonctionnement rhétorique et pragmatique de l'écrit scientifique. En nous fondant sur le modèle du *Corpus Pattern Analysis*, nous élaborons une ressource lexicale des patrons verbaux à l'aide de techniques du TAL sur des corpus numériques. Notre objectif est d'intégrer à cet outil les collocations et les constructions verbales récurrentes dans les écrits scientifiques. Cette ressource a pour objectif d'aider les locuteurs allophones dans leur rédaction d'écrits académiques en français.

Sur le plan didactique, afin de mieux comprendre les difficultés des étudiants allophones en matière de constructions verbales scientifiques, nous établissons un diagnostic qui comporte à la fois l'analyse des erreurs et celle des cas de suremploi et de sous-emploi de certaines de ces constructions verbales. Les résultats ont montré que le statut de novice, en matière de rédaction scientifique, confronte de manière identique les étudiants natifs et non-natifs aux difficultés liées au genre du discours. Afin de répondre à ces besoins langagiers, nous avons proposé d'intégrer la ressource lexicale des patrons dans l'enseignement de l'écrit académique dont l'objectif est de guider les étudiants vers l'autonomie et la manipulation de corpus numériques.

**MOTS-CLÉS : constructions verbales, lexique scientifique transdisciplinaire, linguistique de corpus, patrons, écrits scientifiques**

## ABSTRACT :

The present study focuses on scientific verb constructions, i.e., the lexicon situated between general language and terminology and which is used for scientific expressions, whose rhetorical and pragmatic function must be mastered while writing scientific discourse. Based on the *Corpus Pattern Analysis* model, we develop a lexical resource of verb patterns using NLP techniques, which include collocations and recurring verbal constructions. The purpose of this resource is to help non-native speakers to write their academic writings in French.

From an educational point of view, in order to understand better the difficulties of foreign students with regard to scientific verb constructions, we establish a diagnosis which includes the analysis of both errors and cases of overuse and underuse (notably compared to native students). The results show that both unexperienced natives and non-natives have to face difficulties related to discourse genre. In order to answer these linguistic needs, we suggest that a verb pattern resource should be integrated into corpus-based teaching of academic writing so as to help students gain autonomy in corpus manipulation.

**KEY WORDS : verb constructions, interdisciplinary scientific words, corpus linguistics, patterns, scientific writings**

**TITLE : Study of scientific verbal constructions from a educational perspective : using the corpus in the diagnosis of the language needs of the FLE (French as a Foreign Language) with the techniques of NLP**



---

## REMERCIEMENTS

Je tiens à remercier ici toutes les personnes qui m'ont aidée à réaliser les recherches que j'ai menées durant ces cinq belles années passées à l'Université Grenoble Alpes.

Dans un premier temps, je souhaite remercier deux personnes qui ont scrupuleusement encadré mon travail, ma directrice de thèse Agnès Tutin et ma co-directrice de thèse Cristelle Cavalla, grâce auxquelles j'ai pu associer la linguistique et la didactique. Le savoir scientifique et les qualités de chercheur-enseignant qu'elles ont déployés, ont aidé à former mon esprit à la science et vont influencer considérablement ma carrière professionnelle.

J'adresse mes remerciements les plus sincères aux membres du jury pour l'intérêt qu'ils ont bien voulu manifester quant au sujet de la thèse. Je remercie Natalie Kübler et Jan Goes d'avoir accepté d'être les rapporteurs de cette thèse, ainsi que John Osborne et Francis Grossmann d'avoir accepté de faire partie du jury.

Je remercie toute l'équipe du laboratoire Lidilem, au sein duquel j'ai réalisé ce travail. Un merci tout particulier à Sylvain Hatier pour nos collaborations. Je voudrais remercier chaleureusement Thi Thu Hoai Tran pour nos nombreux échanges scientifiques et amicaux, Marie-Paule Jacques pour sa formation sur le logiciel TXM, Achille Falaise pour sa formation sur l'outil Dicorpus, Magdalena Augustin pour son aide dans le travail du classement des verbes. Un grand merci à Zohra Bouhania pour ses sourires et ses aides dans les démarches administratives.

Je tiens aussi à remercier Mariarosaria Gianninoto de m'avoir accordé sa confiance, ses conseils et ses encouragements répétés tout au long de mes enseignements et mes recherches au Lansad. Ceux-ci ont permis de financer ma thèse. Merci à tous mes collègues du projet Innovalangues : Yoann Goudin, Wang Jinjing, Yuan Zhongjun, Zhang Yijie, Chen Pinxue, Mathieu Loiseau, Julie Attard, Cervini Cristina, Pauline Ballot, Racha Hallal, et tant d'autres, dont la compagnie en contexte professionnel a été réellement enrichissante.

---

Au-delà de mon université, mes remerciements vont au groupe de recherche RIILP de l'Université Wolverhampton qui m'a accueillie pour un séjour de recherche. Je tiens à remercier Patrick Hanks et Sara Moze pour m'avoir aidée à mieux comprendre la construction des patrons CPA dans le PDEV.

Mes remerciements s'adressent également aux membres du département de français de l'Université des Études Internationales de Xi'an qui m'ont accordé leur confiance pour un stage d'enseignement de français. Je remercie tout spécialement JIA Baojun et ZHANG Hongjie qui m'ont aidée à mener à bien mon expérimentation didactique.

Merci également à tous les étudiants qui ont accepté de participer à mes expérimentations didactiques.

Les soutiens des amis et de la famille sont essentiels pour la vie. J'ai eu la chance d'avoir des amis qui m'ont entourée pendant la période de rédaction. Merci à mes deux amis français connectés sur QQ, Maryvonne et Jean, pour nos échanges tant culturels que linguistiques, leur disponibilité et leur infatigable relecture de la thèse. Je remercie également Yoann et Fabrice d'avoir corrigé mon travail même s'ils sont eux aussi en période de finalisation de la thèse. Merci à Fanny, Anita, He, Liwen, Tingting et Yanru, qui m'ont presque permis d'oublier ma solitude.

Avec le recul, ces années n'ont pas été faciles pour mes parents et pour mes beaux-parents. C'est pourquoi je les remercie énormément de m'avoir permis de vivre ma vie, me démontrant ainsi, s'il en était besoin, toute l'étendue de l'amour qu'ils me portent. Un merci tout particulier à ma maman qui s'est montrée tellement compréhensive pour me soutenir.

J'ai aussi, bien sûr, une pensée toute particulière pour l'homme de ma vie, Lei, qui m'a accompagnée pendant tous les hauts et les bas de la vie et n'a jamais douté de mes compétences.

Je dédie cette thèse à mes grands-parents qui m'ont quittée pendant le cours de mes études. J'aurais tellement aimé leur rendre un peu de tout ce qu'ils m'ont prodigué, et qui a fait ma force, en leur faisant partager ma fierté et mon bonheur d'avoir réussi, car c'est à eux surtout que je dois d'être celle que je suis aujourd'hui !





---

# SOMMAIRE

<b>LISTE DES ILLUSTRATIONS.....</b>	<b>XII</b>
<b>INDEX DES TABLES .....</b>	<b>XIV</b>
<b>INTRODUCTION.....</b>	<b>1</b>
1. CONTEXTE DE L'ETUDE.....	1
2. INTERET ET OBJECTIFS DU TRAVAIL .....	2
3. LA LINGUISTIQUE DE CORPUS ET LA LEXICOGRAPHIE ASSOCIEES AU TRAITEMENT AUTOMATIQUE DES LANGUES (TAL) AU SERVICE DE LA DIDACTIQUE DE L'ECRIT .....	3
4. PLAN DE LA THESE .....	5
5. HYPOTHESES ET QUESTIONS .....	6
<b>CHAPITRE I. LES PATRONS VERBAUX DU LEXIQUE SCIENTIFIQUE</b>	
<b>TRANSDISCIPLINAIRE.....</b>	<b>8</b>
1.1. LES PATRONS LEXICO-SYNTAXIQUES DES VERBES : DES ELEMENTS PRECONSTRUITS ASSOCIANT UNE CONSTRUCTION SYNTAXIQUE ET UN SENS .....	9
1.1.1. Verbes et constructions dans la linguistique française.....	9
1.1.2. La notion de construction dans la linguistique cognitive (Fillmore, Goldberg) .....	17
1.1.3. Des constructions aux patrons dans la linguistique contextualiste anglo-saxonne .....	21
1.2. LES ECRITS SCIENTIFIQUES ET LES VERBES DU LEXIQUE SCIENTIFIQUE TRANSDISCIPLINAIRE	28
1.2.1. Écrit scientifique et écrit académique.....	29
1.2.2. Le lexique scientifique transdisciplinaire (LST) et le projet Scientext...	30
1.2.3. Verbes dans <i>l'English for Academic Purposes</i> .....	33
1.2.4. Les patrons des verbes du LST dans le cadre du CPA .....	39
1.3. LA PHRASEOLOGIE DU LEXIQUE SCIENTIFIQUE TRANSDISCIPLINAIRE : DEFICITS PERSISTANTS DES APPRENANTS NON-NATIFS .....	42
1.4. PROBLEMATISATION ET REFLEXION EN VUE DES APPLICATIONS PEDAGOGIQUES SUR LES PATRONS VERBAUX .....	44
1.4.1. Combiner l'approche syntactico-sémantique et onomasiologique : le modèle de Pecman et son application dans <i>l'Aide à la Rédaction de TExtes</i>	

---

<i>Scientifiques (ARTES)</i> .....	45
1.4.2. Intégrer les corpus dans l'élaboration et l'utilisation du dictionnaire : <i>The Louvain EAP Dictionary (LEAD)</i> .....	49
1.4.3. Appliquer le CPA à l'écrit scientifique : <i>E-Advanced Learner's Dictionary     of Verbes in Science (DicSci)</i> .....	52
1.5. BILAN .....	57
<b>CHAPITRE II. CONSTITUTION ET EXPLORATION DES CORPUS</b> .....	<b>59</b>
2.1. UNE APPROCHE INSCRITE DANS LA LINGUISTIQUE DE CORPUS .....	61
2.1.1. La linguistique de corpus : apports et limites .....	62
2.1.2. Une approche associant les analyses « corpus-based » et « corpus-driven » .....	63
2.2. L'UTILISATION DE TROIS CORPUS DANS NOTRE ANALYSE LINGUISTIQUE AVEC UN OBJECTIF DIDACTIQUE PRECIS .....	64
2.2.1. Les usages de trois corpus dans notre étude .....	64
2.2.2. L'intérêt des corpus dans l'enseignement et l'apprentissage des langues étrangères .....	66
2.3. LES CORPUS D'ANALYSE .....	68
2.3.1. Le corpus d'analyse : Scientext-SHS .....	68
2.3.2. Le corpus d'apprenants : corpus Sup-Chinois .....	70
2.3.3. Le corpus d'étudiants natifs : corpus Sup-Natifs .....	76
2.3.4. Les choix des méthodes d'analyse pour le diagnostic des besoins langagiers .....	78
2.3.5. Le tableau récapitulatif de comparabilité .....	80
2.4. LES OUTILS D'EXPLOITATION .....	81
2.4.1. ScienQuest .....	82
2.4.2. Le Lexicoscope .....	86
2.4.3. TXM .....	89
2.5. BILAN .....	91
<b>CHAPITRE III. LES PATRONS LEXICO-SYNTAXIQUES DES VERBES DU LST : DE L'EXTRACTION A LA MODELISATION</b> .....	<b>92</b>
3.1. EXTRACTION DES VERBES DU LST ET DES CADRES DE SOUS-CATEGORISATION .....	93
3.1.1. Extraction des verbes du LST .....	93

---

3.1.2. Extraction des cadres de sous-catégorisation par TAL.....	94
3.2. TRAITEMENT SEMANTIQUE DES VERBES DU LST.....	97
3.2.1. Repérage des acceptions avec la ressource LVF.....	98
3.2.2. Classification sémantique.....	104
3.3. UTILISATION DU MODELE CPA DANS L'ECRIT SCIENTIFIQUE.....	108
3.3.1. Attribution des types/rôles sémantiques.....	109
3.3.2. Vers des patrons syntaxiques spécifiques au genre.....	114
3.3.3. Représentation des patrons par entrée onomasiologique.....	120
3.3.4. Récapitulatif des points communs et des différences.....	122
3.4. BILAN.....	123
<b>CHAPITRE IV. ANALYSE LINGUISTIQUE DES VERBES DU LST.....</b>	<b>125</b>
4.1. INTERET POUR LES VERBES CHOISIS.....	126
4.1.1. Choix des verbes d'analyse.....	126
4.1.2. Choix des verbes de positionnement.....	126
4.2. PARAMETRES DE L'ANALYSE.....	127
4.3. LES VERBES D'ANALYSE.....	128
4.3.1. Verbes de comparaison.....	129
4.3.2. Verbes de constat.....	136
4.3.3. Verbes de description.....	143
4.3.4. Verbes d'examen.....	148
4.3.5. Verbes d'interprétation.....	154
4.3.6. Récapitulatif.....	160
4.4. LES VERBES DE POSITIONNEMENT.....	163
4.4.1. Verbes d'opinion.....	165
4.4.2. Verbes de démonstration.....	170
4.4.3. Verbes de mise en relief.....	176
4.4.4. Verbes de choix.....	182
4.4.5. Verbes d'hypothèse.....	188
4.4.6. Verbes de proposition.....	194
4.4.7. Verbes d'inférence.....	201
4.4.8. Récapitulatif.....	206
4.5. BILAN.....	209

---

## CHAPITRE V. DIAGNOSTIC DES BESOINS LANGAGIERS CHEZ LES APPRENANTS

<b>CHINOIS.....</b>	<b>211</b>
5.1. TROIS NOTIONS ESSENTIELLES POUR LE DIAGNOSTIC : LES NORMES, LES ERREURS ET LES SPECIFICITES .....	213
5.1.1. Les normes linguistiques propres aux écrits scientifiques .....	213
5.1.2. Les erreurs.....	214
5.1.3. Les spécificités.....	215
5.2. DIAGNOSTIC I : ERREURS RELATIVES AUX VERBES DU LST .....	216
5.2.1. Typologie des erreurs en FLM et FLE.....	216
5.2.2. Constitution d'une typologie des erreurs centrée sur les verbes du LST .....	222
5.2.3. Analyse des erreurs .....	225
5.3. DIAGNOSTIC II : SPECIFICITES LIEES AUX CONSTRUCTIONS VERBALES DU LST .....	240
5.3.1. Études représentatives sur les spécificités dans le domaine de l'EAP..	240
5.3.2. Constitution de la grille de diagnostic sur les spécificités .....	247
5.3.3. Analyse des spécificités.....	250
5.4. RECAPITULATIF DU DIAGNOSTIC.....	271
5.5. PISTES DIDACTIQUES .....	273
5.5.1. Enseignement des patrons avec le corpus .....	274
5.5.2. Deux expérimentations initiales .....	277
5.5.3. Exemples d'exercices adaptés .....	286
<b>CONCLUSION GENERALE .....</b>	<b>289</b>
Résultats essentiels.....	289
Limites et perspectives.....	291
<b>BIBLIOGRAPHIE .....</b>	<b>293</b>
<b>ANNEXE .....</b>	<b>309</b>



---

## LISTE DES ILLUSTRATIONS

Figure 1-1 : Les verbes du LST et les patrons verbaux dans une conclusion d'article d'économie	41
Figure 1-2 : Extrait du schéma collocationnel pour la notion  DESCRIPTION  en français (Pecman, 2004, p. 357)	47
Figure 1-3 : Accès aux unités phraséologiques par fonctions discursives dans la base ARTES	48
Figure 1-4 : Entrée du nom source dans le LEAD (Paquot, 2012, p. 175)	50
Figure 1-5 : Note d'erreurs pour l'entrée de l'expression on the contrary dans le LEAD (ibid., p. 173)	51
Figure 1-6 : Extrait du réseau collocationnel du verbe show (Williams & Millon, 2010, p. 1256)	54
Figure 1-7 : Patrons et collocations du verbe show dans le DicSi (Williams & Millon, 2010, p. 1255)	56
Figure 2-1 : Les usages des trois corpus constitués dans notre étude	65
Figure 2-2 : Sélection d'un sous-corpus dans ScienQuest pour le corpus français de Scientext : articles de recherche dans 10 disciplines SHS	83
Figure 2-3 : Requête pour le verbe observer ayant comme sujet les pronoms nous/on/je et objet les noms	84
Figure 2-4 : Visualisation en lemme des constructions verbales nous/on/je + observer + noms	85
Figure 2-5 : Visualisation de répartition par disciplines	85
Figure 2-6 : Requête pour les co-occurents d'adverbes du verbe analyser	86
Figure 2-7 : Visualisation des co-occurents d'adverbes du verbe analyser	87
Figure 2-8 : Requête pour le verbe analyser ayant comme sujet les pronoms et objet les noms	88
Figure 2-9 : Lignes de concordances pour le verbe analyser ayant comme sujet les pronoms et objet les noms	88
Figure 2-10 : Requête des constructions en comme avec les verbes montrer, indiquer, souligner et expliquer dans le corpus Sup-Natifs	90
Figure 3-1 : Exemple d'analyse en dépendances	95
Figure 3-2 : Les patrons du verbe explain dans le PDEV	120
Figure 3-3 : Représentation des patrons du verbe expliquer avec l'entrée par classe	121
Figure 3-4 : Représentation des patrons et des constructions verbales avec l'entrée par fonction rhétorique-exemple du verbe expliquer	122
Figure 4-1 : La classification des verbes d'analyse à étudier	129
Figure 4-2 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de comparaison)	132
Figure 4-3 : Les noms objets des verbes de comparaison (en %)	134
Figure 4-4 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de constat)	139
Figure 4-5 : Les noms objets des verbes de constat (en %)	141
Figure 4-6 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de description)	145
Figure 4-7 : Les noms objets des verbes de description (en %)	147
Figure 4-8 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes d'examen)	151
Figure 4-9 : Les noms objets des verbes d'examen (en %)	153
Figure 4-10 : Les fréquences relatives des pronoms sujets je, on et nous (en % en	

co-occurrences des verbes d'interprétation)	157
Figure 4-11 : Les noms objets des verbes d'interprétation (en %)	159
Figure 4-12 : Les principales constructions verbales des verbes d'analyse et leur fonctions	162
Figure 4-13 : Classification des verbes de positionnement à étudier	164
Figure 4-14 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes d'opinion)	168
Figure 4-15 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de démonstration)	172
Figure 4-16 : Les sujets métonymiques des verbes de démonstration (en %)	174
Figure 4-17 : Les noms objets des verbes de démonstration (en %)	174
Figure 4-18 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de mise en relief)	179
Figure 4-19 : Les noms objets des verbes de mise en relief (en %)	181
Figure 4-20 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de choix)	185
Figure 4-21 : Les noms objets des verbes de choix (en %)	186
Tableau 4.12 : Patrons lexico-syntaxiques des verbes de choix	188
Figure 4-22 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes d'hypothèse)	191
Figure 4-23 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes de proposition)	197
Figure 4-24 : Les sujets métonymiques des verbes de proposition (en %)	198
Figure 4-25 : Les noms objets des verbes de proposition (en %)	199
Figure 4-26 : Les fréquences relatives des pronoms sujets je, on et nous (en % en co-occurrences des verbes d'inférence)	203
Graphique 4-27 : Les noms objets du verbe déduire (en %)	205
Figure 4-28 : Les principales constructions verbales des verbes de positionnement	208
Figure 5-1 : Répartition des erreurs selon les classes d'erreurs de la typologie	226
Figure 5-2 : Répartition des erreurs selon les types d'erreurs de la typologie	227
Figure 5-3 : Ratio de fréquence relative pour les 20 verbes les plus fréquents de notre liste du LST	252
Figure 5-4 : Ratio de fréquence relative pour les 20 verbes du LST les plus utilisés par les apprenants chinois	254
Figure 5-5 : Ratio de fréquence relative pour les 20 verbes du LST les plus utilisés par les novices français	256
Figure 5-6 : La proportion des constructions Sujet humain + Verbe + que (considérer, etc.)	258
Figure 5-7 : La proportion des constructions sujet humain + choisir + objet nominal (choisir, etc.)	260
Figure 5-8 : La proportion des constructions passives	263
Figure 5-9 : La proportion des constructions sous-employées ou absentes chez les apprenants chinois	265
Figure 5-10 : La proportion des constructions suremployées chez les apprenants chinois	268
Figure 5-11 : Entrée onomasiologique de l'outil Dicorpus 'faire un rappel du travail présenté'	281
Figure 5-12 : L'activité rédactionnelle sur l'outil Scribplus	282

---

## INDEX DES TABLES

Tableau 1.1 : Les constructions verbales du verbe dire dans le Lexique-Grammaire	11
Tableau 1.2 : Comparaison des constructions verbales du verbe dire dans le Lexique-Grammaire et dans l'étude de Willems	14
Tableau 1.3 : Les 13 emplois du verbe dire répertoriés dans le LVF	17
Tableau 2.1 : Composition du corpus d'analyse	69
Tableau 2.2 : Récapitulatif de la comparabilité de trois corpus d'analyse	81
Tableau 3.1 : Trois cadres de sous-catégorisation du verbe montrer	96
Tableau 3.2 : Repérage des acceptions des sens relevant de l'écrit scientifique à partir du modèle lexicographique LVF	98
Tableau 3.3 : L'entrée 10 du verbe appliquer dans le LVF	101
Tableau 3.4 : L'entrée 02 du verbe postuler dans le LVF	102
Tableau 3.5 : L'entrée 02 du verbe caractériser dans le LVF	103
Tableau 3.6 : Les cinq entrées du verbe interroger dans le LVF	104
Tableau 3.7 : Acceptions du verbe expliquer dans le corpus Scientext-SHS	106
Tableau 3.8 : Exemples de sous-classes	107
Tableau 3.9 : Les types sémantiques utilisés dans l'analyse des patrons	111
Tableau 3.10 : Comparaison entre le modèle CPA et notre adaptation	123
Tableau 4.1 : Patrons lexico-syntactiques des verbes de comparaison	136
Tableau 4.2 : Récapitulatif des constructions syntaxiques récurrentes des verbes de constat	139
Tableau 4.3 : Patrons lexico-syntactiques des verbes de constat	143
Tableau 4.4 : Patrons lexico-syntactiques des verbes de description	148
Tableau 4.5 : Récapitulatif des constructions syntaxiques récurrentes des verbes d'examen et leurs fonctions correspondantes	151
Tableau 4.6 : Patrons lexico-syntactiques des verbes de description	154
Tableau 4.7 : Patrons lexico-syntactiques des verbes d'interprétation	160
Tableau 4.8 : Récapitulatif des constructions syntaxiques récurrentes des verbes d'opinion et leurs fonctions correspondantes	167
Tableau 4.9 : Patrons lexico-syntactiques des verbes d'opinion	170
Tableau 4.10 : Patrons lexico-syntactiques des verbes de démonstration	176
Tableau 4.11 : Patrons lexico-syntactiques des verbes de mise en relief	182
Tableau 4.13 : Patrons lexico-syntactiques des verbes d'hypothèse	193
Tableau 4.14 : Patrons lexico-syntactiques des verbes de proposition	201
Tableau 4.15 : Patrons lexico-syntactiques des verbes d'inférence	206
Tableau 5.1 : Les types d'erreurs de combinatoire restreinte (Anctil, 2010, p. 194)	218
Tableau 5.2 : Classement des erreurs proposé par Ying (2013, p. 94)	220
Tableau 5.3 : Typologie des erreurs centrée sur les verbes du LST	224
Tableau 5.4 : Grille de diagnostic sur les spécificités	249
Tableau 5.5 : Les fréquences absolue/relative et le ratio de fréquence relative pour le verbe montrer	251
Tableau 5.6 : Récapitulatif de diagnostic	272
Tableau 5.7 : Concordancier du verbe montrer	287
Tableau 5.8 : Exercice de classement sémantique	288
Tableau 5.9 : Correction de l'exercice	288

---

---

# INTRODUCTION

Cette thèse doctorale s'inscrit dans le cadre de deux projets de recherche ANR menés au Lidilem, à savoir les projets Scientext<sup>1</sup> et TermITH<sup>2</sup>. Le projet Scientext (2007-2010) a permis de constituer un corpus d'écrits scientifiques permettant d'effectuer des études linguistiques. Il offre aux chercheurs intéressés par le texte scientifique des ressources textuelles qui leur permettent d'effectuer plusieurs types de recherches dont l'étude de la langue scientifique pour l'aide à la rédaction des écrits universitaires auprès des étudiants étrangers, ce qui constitue le cadre de notre travail. Le projet TermITH (2012-2016) s'est intéressé à l'indexation automatique des textes scientifiques en sciences humaines et sociales. Les travaux menés au laboratoire Lidilem dans le cadre de ce projet avaient pour objectif principal de construire une ressource intégrant les éléments du lexique scientifique transdisciplinaire (LST), ce qui suppose la détection, la désambiguïsation et l'analyse des mots du LST dans les textes. Dans ce cadre, notre travail a consisté à établir les classes des verbes du LST et à identifier les constructions verbales ainsi que les routines sémantico-rhétoriques intégrant des verbes du LST.

## 1. Contexte de l'étude

Selon les chiffres publiés<sup>3</sup> par Campus France, la France se positionne aujourd'hui au 3<sup>e</sup> rang des destinations des étudiants chinois en mobilité, pour plusieurs raisons : l'intérêt culturel du pays, l'apprentissage d'une langue en

---

<sup>1</sup> Projet ANR Agence Nationale pour la Recherche (Gouvernement Français) coordonné par le LIDILEM, ce projet est sous la responsabilité d'Agnès Tutin et Francis Grossmann du LIDILEM. Site du projet Scientext : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article1> (consulté le 23/05/2016)

<sup>2</sup> TermITH (Terminologie et Indexation de Textes en sciences Humaines) : ANR-12-CORD-0029 CONTINT. ATILF, INIST, LIDILEM, LINA, INRIA NGE et Saclay : <http://www.atilf.fr/ressources/termith/> (consulté le 23/02/2016)

<sup>3</sup> Le dossier « Chiffres clés des étudiants internationaux » de Campus France, février 2017 : [http://ressources.campusfrance.org/publi\\_institu/etude\\_prospect/chiffres\\_cles/fr/chiffres\\_cles.pdf](http://ressources.campusfrance.org/publi_institu/etude_prospect/chiffres_cles/fr/chiffres_cles.pdf) (consulté le 23/05/2016)

immersion ainsi que la qualité de la formation universitaire. En 2016, on dénombre 28043 étudiants chinois dont 16 629 (59,3%) étaient inscrits dans une université, soit la troisième communauté derrière les étudiants marocains et algériens en France.

Cependant, ces étudiants chinois, après des années de scolarisation dans le système éducatif chinois, n'ont souvent pas les acquis correspondant aux objectifs académiques « à la française » (Omer, 2013). Or, ils doivent accomplir les mêmes tâches que leurs collègues français natifs pour réussir leurs études en France, souvent des productions écrites, sous la forme de synthèses et de mémoires. Nous nous intéressons particulièrement aux étudiants chinois de 2<sup>e</sup> ou 3<sup>e</sup> cycle dans les universités françaises ou francophones qui ont à rédiger un écrit scientifique (rapport de recherche, mémoire, thèse ou article de recherche). Même si la plupart des étudiants ont atteint un niveau B1/B2 en français, ils sont souvent confrontés à des difficultés rédactionnelles, en particulier celles qui sont liées à la combinatoire lexicale des mots en contexte. Ces phénomènes linguistiques se révèlent être l'une des principales sources des nombreuses erreurs chez les apprenants de Français Langue Etrangère (FLE) (Polguère, 2000). Dans ce travail, nous nous intéressons particulièrement à la phraséologie transdisciplinaire scientifique, et plus précisément aux patrons lexico-syntaxiques des verbes du LST. Il s'agit des mots situés essentiellement à l'intersection entre la langue courante et les langues spécialisées, qui appartiennent à la langue générale, mais sont utilisés à des fins scientifiques.

## **2. Intérêt et objectifs du travail**

L'étude de ces patrons est envisagée dans une conception étendue du domaine de la phraséologie. Nous focaliserons notre attention sur le lien entre un verbe particulier et l'environnement lexico-syntaxique dont le verbe fait partie, car les travaux sur ce sujet montrent que les mots ont tendance à entrer dans un nombre relativement restreint de patrons (Hanks & Pustejovsky, 2005 ; Popescu *et al.*, 2014). Pour un verbe, chaque patron est associé à un sens particulier et obéit à des structures syntagmatiques et à des valeurs sémantiques spécifiques. Ces patrons verbaux sont intégrés dans la phraséologie transdisciplinaire scientifique et les études montrent qu'ils y sont particulièrement surreprésentés (Tutin, 2014 ; Paquot, 2010).

Notre motivation pour les patrons lexico-syntaxiques des verbes du LST est liée aux complexités linguistiques du phénomène (d'ordre morphosyntaxique, syntaxique et sémantique) pour les apprenants et aux exigences de l'écrit scientifique. Ici, nous rejoignons l'avis de Cavalla (2014) selon lequel, l'enseignement et l'apprentissage (ci-après l'E/A) de la phraséologie en général et des patrons lexico-syntaxiques de verbes en particulier pourrait constituer une aide à la lecture et à l'écriture du texte scientifique. En outre, la bonne maîtrise des patrons préfabriqués, dans ce genre, est requise pour écrire un texte scientifique de qualité acceptable pour intégrer la « tribu académique » (Tutin, 2010a).

Dans ce travail, notre objectif est triple.

- a) Nous souhaitons d'abord mener une étude linguistique en montrant les spécificités du comportement lexico-syntaxique des verbes dans ce type d'écrits en examinant plus précisément les verbes de positionnement et les verbes d'analyse.
- b) Ensuite, nous souhaitons établir un diagnostic des besoins langagiers des apprenants chinois en matière de verbes du LST. Les résultats du diagnostic permettront de mieux orienter notre réflexion au plan didactique.
- c) À l'issue de l'étude des patrons et de l'analyse des difficultés langagières, nous réfléchirons à la construction d'une ressource lexicale destinée aux enseignants du FLE. Celle-ci doit mettre en évidence les spécificités du comportement lexico-syntaxique des verbes dans l'écrit scientifique et pourra alors être intégrée dans un outil d'aide à la rédaction scientifique.

### **3. La linguistique de corpus et la lexicographie associées au traitement automatique des langues (TAL) au service de la didactique de l'écrit**

Notre travail est à la croisée de plusieurs disciplines. Mentionnons d'abord, le couple « linguistique de corpus » et « phraséologie ». Depuis son apparition, la linguistique de corpus se focalise sur le phénomène de la phraséologie, des mots

composés aux syntagmes plus ou moins idiomatiques en passant par les locutions, les expressions, les collocations et les patrons. Dans notre étude, la linguistique de corpus permet d'observer les comportements sémantique et syntaxique des verbes, notamment les constructions syntaxiques et les collocations, et d'identifier les patrons lexico-syntaxiques des verbes du LST récurrents dans l'écrit scientifique. Il faut souligner que l'étude de la phraséologie débouche souvent sur des approches fondées sur corpus comme les *pattern grammars* de Hunston et Francis (2000) et le modèle *Corpus Pattern Analysis* (CPA) de Hanks (2013). Ces approches font apparaître des projets lexicographiques comme *Collins COBUILD Grammar Patterns 1 : Verbs* (Hunston & Francis, 1996), *Collins COBUILD Grammar Patterns 2 : Nouns and Adjectives* (Hunston & Francis, 1997) et *Pattern Dictionary of English Verbs* (PDEV) (Hanks, 2013).

Le deuxième volet de ce travail concerne les liens entre la linguistique de corpus et la lexicographie. En partant du modèle d'analyse CPA et en nous basant sur le dictionnaire *Les verbes français* (LVF) de Dubois et Dubois-Charlier (1997), nous souhaitons établir une méthodologie pour créer un prototype de ressource spécialisée des verbes du LST comme outil d'aide à la rédaction scientifique. Comme nous le verrons dans le chapitre I, le besoin d'un dictionnaire spécialisé pour faciliter la production des apprenants est bien réel. Nous pouvons citer quelques outils en cours de construction comme le *Dictionary of Verbs in Science* (*DicSci*) ou le *Louvain English Academic Purposes Dictionary* (*LEAD*) (Granger & Paquot, 2010) pour l'anglais scientifique. En français, il existe deux dictionnaires en ligne, développés à partir de corpus numériques, destinés aux apprenants du FLE : le *Dictionnaire d'apprentissage du français langue étrangère ou seconde* (DAFLES) (Binon et al., 2002) et le dictionnaire d'*Aide à la Rédaction de TExtes Scientifiques* (ARTES) (Pecman & Kübler, 2011). Notre étude vise à combler en partie le manque de ressources spécialisées en ligne dans le domaine du FLE.

Le troisième volet de ce travail se rapporte aux liens entre la linguistique de corpus et la didactique. Bon nombre de chercheurs en didactique affirment que les développements de la linguistique de corpus ont contribué au renouvellement de la réflexion sur les contenus d'enseignement et sur la place des connaissances linguistiques dans l'enseignement (Pecman, 2005). Les trois corpus constitués dans



notre étude ont des visées didactiques précises. D'une part, les corpus d'écrits universitaires d'étudiants (le Sup-Chinois et le Sup-Natifs) permettent d'observer l'utilisation des verbes du LST dans les écrits des étudiants et d'identifier les difficultés des apprenants chinois en amont de la recherche orientée vers la didactique de l'écrit académique. D'autre part, l'utilisation d'un corpus numérique pour la classe de langue présente un intérêt certain pour l'enseignement de la phraséologie. Nous montrons dans le chapitre V comment exploiter les données du corpus TermITH-SHS en vue d'enseigner les constructions verbales des verbes du LST.

#### **4. Plan de la thèse**

Le plan de la thèse s'articule de la façon suivante. Le chapitre I est destiné à exposer notre objet d'étude – les patrons lexico-syntaxiques des verbes du LST –, et à présenter son terrain d'application, l'enseignement de l'écrit académique. Nous précisons d'abord la définition et la description des objets analysés : en premier lieu les constructions puis les patrons des verbes. Nous entrons ensuite dans des définitions plus précises quant à notre objet d'étude ; en d'autres termes, nous expliquons ce que nous entendons par patrons verbaux dans l'écrit scientifique avant de souligner l'importance de leur enseignement en vue d'aider à la rédaction universitaire des étudiants allophones.

Dans le chapitre II, nous présentons la constitution des données utilisées dans l'analyse des verbes et le diagnostic. Nous explicitons les différents usages de trois corpus, mettant en avant les rôles de ces corpus dans cette recherche.

Le chapitre III présente la méthodologie adoptée. Nous décrivons d'abord les méthodes d'extraction des verbes et des cadres de sous-catégorisation à l'aide des techniques de TAL. Nous présentons par la suite les étapes de la modélisation des patrons lexico-syntaxique des verbes du LST en adaptant la méthode CPA à l'étude de l'écrit scientifique. L'unité linguistique choisie est le verbe, catégorie fondamentale de l'énoncé. Dans l'écrit scientifique, le verbe avec son environnement lexico-syntaxique sert à organiser le discours, présenter les activités scientifiques, se positionner, etc. Nous avons choisi de travailler sur deux groupes de verbes : les

verbes d'analyse et les verbes de positionnement, qui sont décrits dans le chapitre IV.

Dans le chapitre V, nous mettons en place les grilles de diagnostic des besoins langagiers. Nous nous appuyons sur les corpus d'écrits universitaires d'étudiants (le Sup-Chinois et le Sup-Natifs) en vue de cerner les erreurs et les spécificités liées à l'utilisation des verbes et des constructions verbales du LST chez les apprenants chinois. Nous proposons également des pistes didactiques dans lesquelles nous décrivons deux expérimentations effectuées auprès des étudiants chinois, ce qui nous conduit ensuite à des réflexions didactiques sur l'enseignement des verbes et des constructions verbales du LST avec l'utilisation des corpus en classe de FLE.

## 5. Hypothèses et questions

Sur le plan linguistique, on peut supposer que les verbes du LST, plus précisément les verbes d'analyse et les verbes de positionnement, présentent des spécificités tant au niveau sémantique que syntaxique dans l'écrit scientifique. Nous postulons également que l'étude des patrons verbaux permet de mettre en évidence ces spécificités. Différentes questions sont ainsi posées : Quel modèle à adopter pour analyser les verbes du LST dans notre étude linguistique à visée didactique ? Quelles données permettent d'observer le comportement lexico-syntaxique des verbes dans l'écrit scientifique ? Quelles sont les constructions verbales fréquemment mobilisées par les verbes d'analyse et les verbes de positionnement, et quelles sont leurs fonctions spécifiques ?

L'étude linguistique des verbes a pour objectif de mieux outiller l'enseignement de la compréhension et de la production des patrons verbaux dans le discours scientifique. En nous situant dans une perspective d'enseignement et d'apprentissage de l'écrit académique, nous partons du constat que par rapport aux étudiants natifs, les apprenants chinois se trouvent confrontés à plus de difficultés à faire passer leurs idées dans la rédaction. La problématique de la recherche porte d'une part, sur le diagnostic des besoins langagiers et d'autre part, sur la transposition didactique des patrons verbaux. Nous souhaitons alors répondre aux questions suivantes : Les difficultés des apprenants chinois sont-elles seulement liées à leur maîtrise de langue ?

Quelle approche adopter pour l'E/A des patrons ? Quel type d'outil à mettre en place ?

---

# Chapitre I.

## Les patrons verbaux du lexique scientifique transdisciplinaire

---

1.1. LES PATRONS LEXICO-SYNTAXIQUES DES VERBES : DES ELEMENTS PRECONSTRUITS ASSOCIANT UNE CONSTRUCTION SYNTAXIQUE ET UN SENS.....	9
1.1.1. Verbes et constructions dans la linguistique française.....	9
1.1.2. La notion de construction dans la linguistique cognitive (Fillmore, Goldberg).....	17
1.1.3. Des constructions aux patrons dans la linguistique contextualiste anglo-saxonne.....	21
1.2. LES ECRITS SCIENTIFIQUES ET LES VERBES DU LEXIQUE SCIENTIFIQUE TRANSDISCIPLINAIRE.....	28
1.2.1. Écrit scientifique et écrit académique.....	29
1.2.2. Le lexique scientifique transdisciplinaire (LST) et le projet Scientext.....	30
1.2.3. Verbes dans l' <i>English for Academic Purposes</i> .....	33
1.2.4. Les patrons des verbes du LST dans le cadre du CPA.....	39
1.3. LA PHRASEOLOGIE DU LEXIQUE SCIENTIFIQUE TRANSDISCIPLINAIRE : DEFICITS PERSISTANTS DES APPRENANTS NON-NATIFS.....	42
1.4. PROBLEMATISATION ET REFLEXION EN VUE DES APPLICATIONS PEDAGOGIQUES SUR LES PATRONS VERBAUX.....	44
1.4.1. Combiner l'approche syntactico-sémantique et onomasiologique : le modèle de Pecman et son application dans l' <i>Aide à la Rédaction de TExtes Scientifiques</i> (ARTES).....	45
1.4.2. Intégrer les corpus dans l'élaboration et l'utilisation du dictionnaire : <i>The Louvain EAP Dictionary</i> (LEAD).....	49
1.4.3. Appliquer le CPA à l'écrit scientifique : <i>E-Advanced Learner's Dictionary of Verbs in Science</i> (DicSci).....	52
1.5. BILAN.....	57

---

Dans ce chapitre, nous présentons différents modèles portant sur les constructions verbales et les patrons verbaux afin de circonscrire la notion de patron lexico-syntaxique adoptée dans notre étude. Dans un deuxième temps, nous abordons notre domaine d'étude, l'écrit scientifique et académique dans lequel s'intègre notre objet de recherche sur les patrons. Ayant un objectif didactique, nous nous penchons ensuite sur les travaux traitant des patrons dans les écrits d'apprenants avant de finir par une présentation de ressources pédagogiques innovantes. Nous proposerons alors quelques pistes pour la construction d'un outil d'aide à la rédaction de l'écrit académique des étudiants allophones du français.

## **1.1. Les patrons lexico-syntaxiques des verbes : des éléments préconstruits associant une construction syntaxique et un sens**

La phraséologie, définie comme « le domaine qui traite les séquences lexicales perçues comme préconstruites, a très largement élargi ses objets d'études » (Legallois & Tutin, 2013, p. 3) et cette conception étendue peut s'expliquer par la nature lexicalisée de la langue elle-même (Sinclair, 1991). Dans notre étude, nous nous intéressons aux patrons des verbes qui se situent à la limite de la phraséologie, à l'intersection des champs du lexique et de la syntaxe. Nous allons étudier ces éléments préconstruits dans le contexte scientifique pour ensuite les appliquer à l'écrit académique. Dans un premier temps, il nous semble primordial d'aborder la notion de construction, et plus précisément, celle des patrons verbaux dans plusieurs courants linguistiques : la linguistique française, la linguistique cognitive américaine et la linguistique de corpus anglo-saxonne dans laquelle s'inscrit l'approche de notre étude.

### **1.1.1. Verbes et constructions dans la linguistique française**

En linguistique française (qui inclut ici les représentants suisses et belges), des chercheurs ont tenté de relier les formes du langage à leur sens. En 1909, les régularités constructionnelles sont abordées sous l'aspect phraséologique par Bally dans son ouvrage, le *Traité de stylistique française*. Selon lui, la phraséologie peut être lexicale, mais aussi grammaticale. Par exemple, *la maison de Paul*, la construction *le/la N de SN* est associée à la signification de la propriété :

D'une manière générale on pourrait prouver qu'une foule de faits de syntaxe ne sont que des groupements phraséologiques à éléments séparables. (Bally, 1909, p. 76)

#### **1.1.1.1. Le lexique-grammaire**

L'étude des constructions reste toutefois succincte et peu systématique jusqu'à l'apparition du *Lexique-Grammaire*, élaboré sous la direction de Maurice Gross qui vise à mettre en évidence les propriétés distributionnelles et transformationnelles des

verbes, dans la droite ligne de l'approche de Harris, dont il a été le disciple. Il s'inspire du modèle de Harris et de la linguistique distributionnelle. En mettant en contraste des ensembles de constructions verbales identiques, mais associées à différents verbes, Gross constate qu'elles se différencient selon une série de propriétés syntaxiques qui ont « une fonction classifiante qui semble être liée aux propriétés sémantiques du langage » (Gross, 1986, p. 171) : temps, présence d'une complétive *que*, transformation passive, nombre et nature des compléments, types de prépositions (*à, de*), réalisations du sujet et des compléments (nom humain, nom au pluriel, nom concret), etc. Par exemple, deux verbes comme *dire* et *vouloir* partagent la même construction et les mêmes arguments suivants :

- (1) *Il dit avoir de l'argent.*
- (2) *Il veut avoir de l'argent.*

Mais parce que l'exemple (1) peut se transformer en une complétive et un objet indirect (*Il m'a dit qu'il a de l'argent.*) alors que l'exemple (2) aura seulement une complétive (*Il veut qu'il soit riche.*), ces deux verbes relèvent de deux tables<sup>1</sup> de constructions différentes (table 9 et table 6). En d'autres termes, ici, la transformation en complétive précédée d'un objet indirect constitue une propriété définitoire du premier exemple. Suivant ce principe, les constructions syntaxiques de 6000 verbes ont été classées dans 81 tables et représentées par des cadres de sous-catégorisation du type *N0 V Qu P à N2* (table 9). Prenons l'exemple de la table 9. Cette grande classe syntaxique regroupe environ 400 verbes qui peuvent être répartis en une série de sous-classes dont des verbes sémantiquement homogènes et intervenant dans la communication : *affirmer, annoncer, avouer, confirmer, demander, dire, évoquer*, etc. Mais ce travail de classement des verbes sémantiquement homogènes au sein de chaque table ne fait pas l'objet d'une présentation détaillée dans l'ouvrage de Gross. Par ailleurs, il faut noter que cette construction inclut aussi des verbes peu utilisés tels que *bourdonner, bramer, chuintier, tonitruer*, ce qui est dû au faible nombre de données et au critère de fréquence.

---

<sup>1</sup> Chaque table regroupe les verbes ayant des propriétés similaires. Celles-ci sont présentées par un codage en « + » et « - » précisant si le verbe concerné possède ou non cette propriété. Ainsi, chaque table constitue une classe syntaxique. Les tables du *Lexique-Grammaire* sont téléchargeables en ligne : <http://infolingu.univ-mlv.fr/>

Il est à noter que la polysémie des verbes, présentés selon leurs propriétés syntaxiques, n'est pas clairement précisée. Par exemple le verbe *dire*<sup>2</sup> appartient à trois tables différentes qui dénotent chacune un sens, mais ce dernier n'est pas explicité et devra être déduit de l'exemple :

Tables	Constructions syntaxiques	Exemples
9	$N0 V Qu P \grave{a} N2$	Max dit à Luc qu'il est malade.
32NM	$N0 V N1$	Ses yeux disent sa fatigue.
32R3	$N0 V N1$ (N1 = nom contraint)	Max a dit un poème.

**Tableau 1.1 : Les constructions verbales du verbe *dire* dans le Lexique-Grammaire**

Comme vu précédemment, la table 9 est définie par la présence d'une complétive précédée d'un objet indirect en *à*. Concernant la table 32 NM, la propriété définitoire réside dans le fait que la construction  $N0 V N1$  ne peut pas se transformer au passif : *No être V-pp par No (Sa fatigue est dite ?)*. La table 32R3 comporte la même construction que la table 32NM, mais celle-ci est définie par une propriété distributionnelle qui porte sur l'objet direct  $N1$ . Ce dernier, défini comme nom contraint, se limite à une petite classe de noms tels *mot, poème, texte*.

Les tables (classes syntaxiques) des verbes sont constituées au moyen de certaines propriétés syntaxiques qui permettent de dégager plus tard des propriétés sémantiques. Par exemple, les verbes de la table 9 suggèrent une idée de « communication ».

Une analyse ne doit pas rendre compte du sens d'une forme donnée, mais plutôt du sens des formes quasi synonymes, et ceci au moyen de propriétés syntaxiques, voisines. [...] les méthodes d'association de la forme et du sens sont, au mieux, rudimentaires. (Gross, 1975, p. 224)

Cette méthode constitue la première étude systématique sur les constructions verbales en français, et de ce fait, a inspiré la constitution de ressources des constructions verbales telles que *Les verbes français* de Jean Dubois et Françoise Dubois-Charlier (1997) que nous allons présenter plus loin.

<sup>2</sup> Les tables sont mises en ligne par le Laboratoire d'Informatique Gaspard-Monge (LIGM). Consulté le 6 novembre 2015 : <http://infolingu.univ-mlv.fr/DonneesLinguistiques/Lexiques-Grammaires/Présentation.html>

### 1.1.1.2. L'étude des constructions verbales de Willems

À l'instar du *Lexique-Grammaire*, l'étude de Willems (1981) place le lexique, plus précisément le verbe, au centre du classement des constructions verbales. Dans ce modèle, le verbe constitue le noyau pour établir le lien entre syntaxe et sémantique, dans la mesure où l'étude syntaxique est poursuivie jusqu'à obtenir des classes de verbes plus petites, mais syntaxiquement et sémantiquement plus homogènes par rapport au modèle du *Lexique-Grammaire*. La procédure de classement de 3000 verbes mettant en relation la syntaxe et la sémantique consiste à prendre en compte trois aspects principaux :

- (1) Les relations sémantiques générales, qui sélectionnent
- (2) un nombre restreint de constructions syntaxiques
- (3) le choix d'une construction implique le choix d'un certain lexique. (*ibid.*, p. 239)

Reprenons l'exemple du verbe *dire*, il entre dans la construction *N1 (quelqu'un) V à N2 (quelqu'un) que P (Il dit à Marie qu'il va partir.)* en partageant cette construction avec un petit groupe de verbes tels *déclarer, affirmer, indiquer, expliquer, annoncer*. Le sens de déclaration est donc lié à cette construction que l'auteure appelle « construction déclarative » (*ibid.* p. 13). Elle avance que tous les verbes, quel que soit leur sens de base, peuvent devenir des « verbes déclaratifs », dès qu'ils adoptent les propriétés syntaxiques caractéristiques de cette classe verbale (Willems, 1983, p. 426). Dans sa classification, elle a identifié cinq types essentiels de complémentation :

- absence de compléments « complémentation intransitive » (ex. *il pleut, il marche*)
- compléments nominaux « complémentation transitive » (ex. *il ressemble à Paul, le jardin pullule d'abeilles, il donne un livre à Marie*)
- compléments adjectivaux ou substantivaux « complémentation attributive » (ex. *il devient grand, il rend sa mère malade*)
- compléments infinitifs « complémentation infinitive » (ex. *il vient jouer, il cherche à réussir, il envisage de partir demain*)
- compléments phrastiques « complémentation complétive » (ex. *je sais que tu*



*pars*).

Dans le cas où un verbe admet plusieurs constructions, il sera classé en fonction de la hiérarchie suivante : complétive > infinitif > constructions nominales indirectes > constructions nominales directes. Ainsi, en termes de classement, la construction en complétive est plus discriminante que la construction à l'infinitif, et ainsi de suite. Par exemple, pour le groupe de verbes (*aimer, adorer, détester*) ci-après qui admettent à la fois la complétive et l'infinitif, le verbe sera classé dans la construction « complétive » :

- (1) *Il aime/adore/déteste venir.*
- (2) *Il aime/adore/déteste que je vienne.*

D'après Willems (1981), ce groupe est sémantiquement plus proche que le groupe des verbes (ex. *approuver, critiquer*) qui n'admet que la construction complétive. De ce fait, « la valeur sémantique de la construction complétive l'emporte sur celle de l'infinitif. » (*ibid.*, p. 48) La proposition de cette hiérarchie permet d'une part, de prendre en considération les valeurs sémantiques des constructions verbales et d'autre part, de rendre les classes des verbes plus homogènes sémantiquement.

Au niveau sémantique, en ce qui concerne les arguments nominaux, Willems distingue cinq sous-catégories pour affiner le classement : nom de personne (P), nom de chose (C), noms de lieu (L), noms désignant une partie du corps (PC), nom d'action (A).

Enfin, l'analyse des constructions verbales a débouché sur une cinquantaine de classes de verbes syntaxiquement et sémantiquement homogènes (mouvement, état, jugement, appartenance, etc.).

Afin de cerner les différences entre ce modèle et celui du *Lexique-Grammaire*, nous proposons ci-dessous un tableau comparatif qui illustre les constructions verbales du verbe *dire* présentées dans ces deux modèles.

Constructions verbales	Le modèle de Willems	Lexique-Grammaire
Constructions en commun	P1+V+C+à+P2/que donner une information Ex : <i>Il m'a dit une nouvelle. /Il m'a dit qu'il va partir.</i>	N0 V Qu P à N2 Table 9 Ex : <i>Max me dit qu'il est malade.</i>
	P1+V+à+P2+Inf./que déclaration d'opinion Ex : <i>Il me dit avoir été trompé. /Il me dit qu'il a été trompé.</i>	
	P1+V+à+P2+de+Inf. /que_subjonctif déclaration de volonté Ex : <i>Il dit à Jean de venir. /Il dit à Jean qu'il vienne.</i>	
Constructions différentes		N0 V N1 Table 32NM Ex : <i>Ses yeux disent sa fatigue.</i>
		N0 V N1 (N1 = nom contraint) Table 32 R3 Ex : <i>Max a dit un poème.</i>

**Tableau 1.2 : Comparaison des constructions verbales du verbe *dire* dans le *Lexique-Grammaire* et dans l'étude de Willems**

Trois constructions sont distinguées par Willems : la première construction rassemble les verbes tels *annoncer*, *conter*, *raconter*, *confier* partageant le sens de 'donner une information'. La deuxième construction admettant à la fois une complétive *que* et l'infinitif s'associe à un petit groupe de verbes d'opinion comme *prétendre*, *avouer*, *affirmer*, *nier*, *reconnaître*. La troisième construction regroupe des sous-ensembles de verbes sémantiquement homogènes : *conseiller*, *demander*, *commander* et l'infinitif prépositionnel peut être remplacé par une complétive au subjonctif. Si l'on considère le sens dégagé, on constate que la première construction ayant le sens de 'donner une information' peut correspondre à la table 9 du *Lexique-Grammaire* qui suggère le sens de « communication ». Mais les deux dernières peuvent également être intégrées dans la table 9 du *Lexique-Grammaire* puisque qu'il manque une analyse sémantique. En revanche, les constructions appartenant aux tables 32NM (*Ses yeux disent sa fatigue*) et 32R3 (*Max a dit un poème*) sont absentes dans le modèle de Willems, ce qui peut s'expliquer par la hiérarchie des compléments définie par Willems. En effet, les constructions nominales directes, situées au bas de la hiérarchie, sont exclues de l'analyse du verbe si elles peuvent être remplacées par les autres constructions précédentes, ce qui posera un problème lorsque l'emploi des constructions nominales directes entraîne une

différence sémantique importante.

De façon générale, nous pensons que ces deux modèles ont mis en évidence le lien étroit entre les constructions verbales et leur interprétation sémantique. Cependant, dans l'étude de Willems, l'analyse sémantique occupe une place plus importante par rapport au modèle du *Lexique-Grammaire*. À l'issue de l'étude de ces deux modèles, on retient que l'analyse des constructions verbales permet de dégager des caractéristiques syntaxiques et sémantiques du verbe. Il faut souligner que ces deux modèles, n'ayant pas une visée lexicographique, ne couvrent que certains sens du verbe et montrent donc quelques lacunes dans la description des constructions verbales. Nous allons maintenant présenter un modèle lexicographique des verbes français qui nous paraît plus exhaustif.

### 1.1.1.2. Les verbes français (LVF)

Dans la continuité des travaux précédents, surtout le *Lexique-Grammaire* de Gross, J. Dubois et F. Dubois-Charlier proposent la classification des verbes français autour des propriétés sémantiques et syntaxiques. C'est un modèle lexicographique exhaustif des verbes français couvrant « 25 610 entrées verbales simples représentant 12 310 verbes différents dont 4 188 à plusieurs entrées » (Dubois & Dubois-Charlier, 1997, p. 1). Il existe en deux versions (papier et électronique) et est accessible en ligne <sup>3</sup>. Fondé sur les méthodes de la grammaire distributionnelle et transformationnelle, il montre l'adéquation entre les constructions syntaxiques et l'interprétation sémantique par les classes syntactico-sémantiques.

Sur le plan sémantique, la description se fait à l'aide d'une suite de caractères alphabétiques dénommée « opérateur » qui « interprètent sémantiquement les schèmes syntaxiques » (*ibid.*, p. 7). À chaque entrée verbale correspond un opérateur complexe (composé d'opérateurs simples) différent qui contribue à la construction des classes syntactico-sémantiques des verbes. Par exemple, l'opérateur de l'entrée *dire* 2 (un emploi parmi les 13) « [dic] information A qn » veut dire 'communiquer une information à quelqu'un'. Il indique que l'entrée *dire* 2 appartient au sous-type syntaxique C2a.1 qui se lit « dire quelque chose, que (temps indicatif) à quelqu'un ».

<sup>3</sup> Consulté le 23 novembre 2015 : <http://rali.iro.umontreal.ca/rali/?q=fr/LVF>

Chaque lettre ou chiffre indique un niveau de classement et il existe cinq niveaux de classement. Ainsi, ce sous-type syntaxique relève de la sous-classe syntaxique P1g « dire que, dire quelque chose à quelqu'un » qui, à son tour, appartient à la classe sémantico-syntaxique P1 « dire, demander quelque chose » et la classe générale « verbes de communication ».

Sur le plan syntaxique, chaque variante syntaxique est exprimée à l'aide d'un schème de construction syntaxique qui comporte une association de lettres et de chiffres entre crochets (voir la rubrique construction ci-dessous). À chaque entrée verbale peuvent correspondre un ou deux schèmes de construction (dont l'un relève souvent de l'emploi pronominal). Par exemple, le schème de construction de l'entrée *dire* 02 « T14a0 » se lit « transitif direct à sujet humain, et suivi par une complétive ou un complément chose ». Les opérateurs et les schèmes font ressortir les propriétés sémantiques et syntaxiques des verbes et constituent deux éléments primordiaux du LVF. Nous présentons ci-dessous les entrées du verbe *dire* dans le LVF en le comparant aux deux modèles précédents :

mot	construction	opérateur	classe	domaine	sens	phrase
dire 01	T1306	[dic] mots	C2f.1	LAN	prononcer	On d~ "pallier un défaut".
dire 02	T14a0	[dic] information A qn	C2a.1	LIT	communiquer	On d~ à P un secret.
dire 03	T15a8	[dic] ordre A qn DE	C2b.2	LIT	ordonner	On d~ à P de venir, qu'il se rende en cet endroit.
dire 04	T13a8 P30a0	- [dic] mot A qn	C2c	LAN	raconter	On d~ des bêtises à P. Ceci ne peut se d~ à P.
dire 05	T1310 P3000	- [dic] énoncé auprès	C2e.1	LIT	énoncer	On d~ une évidence, une bêtise.
dire 06	T1418 P1510	- [dic] QUE/soi tel	C2d.1	LIT	prétendre	On d~ que P est bon, aimer le café. On se d~ intelligent.
dire 07	T13a8	[dic] récit A qn	C2c	ENS	réciter	On d~ des poèmes à l'assistance.
dire 08	T3300	(qc) [ind] abstrait	C4b.3	PSYt	montrer, manifester	Ses yeux d~ sa fatigue, sa joie.
dire 09	N3a	(qc) [f.sent] bien A qn	P2b	PSY	convenir, plaire à	Cette promenade ne me d~ rien.
dire 10	T33a0	(qc) [ind] sent A qn	C4a	PSY	faire penser, évoquer	Ce visage ne d~ rien à P.
dire 11	T1418	[dic] qc/QUE auprès	C2d.1	LIT	raconter	On d~ que P est l'assassin.
dire 12 (se)	P1500	[ger.mens] QUE	P1g.2	PSY	croire	On se d~ qu'il faut y aller.

dire 13	T1318 P3008	-	([ind] abstrait par qc	C4c.2	LITt	exprimer, décrire	L'écrivain d~ la tristesse dans son poème. Ses vers d~ l'amour.
------------	----------------	---	---------------------------	-------	------	----------------------	--

**Tableau 1.3 : Les 13 emplois du verbe *dire* répertoriés dans le LVF**

Le tableau 1.3 montre bien la relation étroite entre syntaxe (schème de construction) et sémantique (opérateur et sens) qui contribue à la construction des sens assez complets et bien distincts. L'entrée 2 du verbe ayant le sens 'communiquer' et la construction syntaxique « transitif direct à sujet humain, et suivi par une complétive ou un complément chose » est partagée par les deux modèles précédents. Les entrées 7 et 8 peuvent se retrouver dans le modèle du *Lexique-Grammaire* tandis que les entrées 3 et 6 peuvent correspondre respectivement aux constructions qui expriment une déclaration de volonté (P1+V+à+P2+de+Inf./que\_subjonctif) et d'opinion (P1+V+à+P2+Inf./que) dans le modèle de Willems. Il faut noter que certains sens ne sont liés qu'à un domaine particulier comme l'entrée 13 (*L'écrivain dit son stress dans son poème*).

En raison de sa large couverture, non seulement en termes du nombre de verbes français répertoriés, mais aussi des propriétés sémantiques et syntaxiques des verbes, cette ressource qui est une référence linguistique pour l'étude des verbes français sera exploitée pour notre travail sur corpus que nous allons développer dans le chapitre III.

Même si la notion de construction n'est pas encore théorisée dans ces travaux, ils ont très tôt découvert des régularités constructionnelles en mettant en valeur l'interaction entre les propriétés sémantiques et les constructions syntaxiques dans la classification des constructions verbales. En outre, ils proposent des approches similaires en plusieurs points aux modèles des *Grammaires de Construction* que nous allons présenter dans la section suivante, mais bien avant les travaux américains.

### **1.1.2. La notion de construction dans la linguistique cognitive (Fillmore, Goldberg)**

La notion de construction est théorisée dans les modèles des *Grammaires de Construction* par un ensemble de chercheurs associés au courant de la linguistique

cognitive. Fillmore et *al.* (1988) considèrent que les constructions peuvent être grammaticales, mais aussi lexicales comme les expressions idiomatiques (le connecteur *let alone*) dans la mesure où elles sont dotées d'une interprétation sémantique qui est non prédictible à partir des composants. Ce point de vue se rapproche de l'idée de Bally et il est partagé par Goldberg qui revendique le caractère non prédictible des propriétés grammaticales, sémantiques et pragmatiques des constructions :

A construction is [...] a pairing of form with meaning/use such that some aspect of the form or some aspect of the meaning/use is not strictly predicable from the component parts or from other constructions already established to exist in the language (Goldberg, 1998 : 205).

Ces constructions se présentent sous des formes variées (Goldberg, 2003, p. 217) telles que des morphèmes (*anti-*), des mots (*and*), des idiomes (*let alone*), et des structures syntaxiques dont les éléments lexicaux sont partiellement saturés :

The Xer the Yer : *the more you think about it, the less you understand.*  
(traduction de l'auteur : Plus tu y penses, moins tu comprends.)

Dans cette construction, les éléments lexicaux X et Y ne sont pas déterminés et restent ouverts à un paradigme lexical large. Dans d'autres constructions, les éléments lexicaux peuvent aussi être totalement saturés :

*all of a sudden* (traduction de l'auteur : tout à coup)

Ici, les éléments lexicaux sont fixes. Selon Goldberg (*ibid.*, p. 221), ces constructions sont des « stored pairings of form and function » (paires des formes et des fonctions mémorisées) qui établissent le lien entre la forme et l'interprétation sémantique. Par exemple, dans la phrase suivante :

*She gave Paul a book.* « X causes Y to receive Z »

La construction ditransitive (Subj V Obj 1 Obj 2) possède un sens indépendant que Goldberg définit comme « X causes Y to receive Z ». Il s'agit donc d'une action prévue ou réelle de transfert de X à Y. Le « sens » (« contenu », voire « fonction » chez certains) d'une construction, relève non seulement des propriétés sémantiques, mais aussi, le cas échéant, de propriétés qui peuvent être considérées comme

pragmatiques ou discursives. D'après Goldberg, la construction ditransitive impose que l'objet 1 (*Paul*) doit être animé, capable d'assurer l'action du transfert et que l'objet 2 (*book*) renvoie à un thème, ces propriétés relèvent de la signification même de la construction et non au verbe (*give*).

Un autre point soulevé par Goldberg concerne le caractère polysémique de la construction. Par exemple, la construction ditransitive peut aussi impliquer la négation :

*She refused Paul a book.* « X causes Y not to receive Z »

De ce point de vue, comme le soulignent Legallois et Gréa (2007), il est raisonnable de considérer qu'ici, les constructions ont les mêmes propriétés que le lexique (forme, sens et polysémie) :

L'implication la plus fondamentale est alors que la construction considérée devient une forme au sens gestaltiste et phénoménologique du mot — entendons une « structure » signifiante — sans que soit posée de distinction préalable entre syntaxe et sémantique. (Legallois & Grea, 2007, p. 167)

Si l'on considère les modèles examinés dans la linguistique française, dans lesquels les différentes classes sémantiques des verbes peuvent partager la même construction syntaxique, on constate que ceci se rapproche de l'idée de la polysémie de construction. Par contre, dans les modèles français, l'unité polysémique porte sur le verbe et non pas les constructions verbales. Il faut noter que les constructions comme la construction ditransitive (Subj V Obj Obj 2) dont la description sémantique des arguments est absente, sont ouvertes aux sens (les linguistes cognitifs parlent de la productivité des constructions). Ce manque de précision au niveau sémantique peut être dû en grande partie au manque de données.

Depuis des années, certains linguistes cognitifs ont intégré le corpus dans l'analyse des constructions. En appliquant l'approche des *Grammaires de Construction* au français, J. François et M. Sénéchal se basent sur *Frantext* et *Le Monde* (année 1993) pour étudier les verbes de production de parole dans trois cadres prédicatifs dotés d'un même prototype sémantique « interaction entre deux partenaires X et Y respectivement exprimés par le sujet et l'objet (direct ou

prépositionnel) visant l'accomplissement (ou le non accomplissement) de A par Y ou l'accomplissement de A par X » (François & Sénéchal, 2006, p. 67) :

- (1)  $N1_{(X)} V N2_{(Y)} \text{ à } INF_{(A)}$  (*obliger*) : *Il oblige Marie à partir.*
- (2)  $N1_{(X)} V N2_{(Y)} \text{ de } INF_{(A)}$  (*empêcher*) : *Il empêche Marie de partir.*
- (3)  $N1_{(X)} V \text{ à } N2_{(Y)} \text{ de } INF_{(A)}$  (*permettre*) : *Il permet à Marie de partir.*

De façon générale, ces trois cadres prédicatifs étant trois variantes regroupées sous un même prototype sémantique qui portent la signification communicationnelle, chaque cadre est analysé en repérant les verbes associés. Par exemple, le cadre prédicatif  $N1_{(X)} V N2_{(Y)} \text{ à } INF_{(A)}$  est exploité par 28 verbes dont *obliger*, *aider*, *inviter* et *amener* pour citer les plus fréquents. Il s'agit d'accomplir A (l'action de partir dans l'exemple) et Y (Marie dans l'exemple) est en cause. Cette étude a montré que les Grammaires de Construction permettent d'expliquer les cas de polysémie verbale (l'interprétation du verbe est différente selon les constructions dans lesquelles il entre) et le fait qu'un verbe possède plusieurs constructions (Legallois, 2005, p. 112).

En résumé, les principales caractéristiques des constructions dans les *Grammaires de Construction* sont les points suivants :

- les constructions possèdent un caractère préconstruit comme des unités phraséologiques, lexicalement non saturées ;
- elles supportent un sens sémantique indépendant du lexique ;
- elles permettent d'expliquer la polysémie des verbes.

Comme nous l'avons mentionné plus haut, notre étude se fonde sur l'approche des patrons, et plus précisément sur l'approche CPA. Cette approche accorde un rôle important aux constructions verbales et aux collocations pour désambiguïser les sens du verbe. Il nous semble important de justifier notre choix dans la section suivante où nous allons présenter la notion des patrons par rapport à celle des constructions.



### 1.1.3. Des constructions aux patrons dans la linguistique contextualiste anglo-saxonne

#### 1.1.3.1. Les patrons grammaticaux dans la linguistique contextualiste

A.S. Hornby a été précurseur dans son intérêt pour les patrons grammaticaux. En tant que professeur d'anglais langue étrangère, il a très vite découvert que les apprenants d'anglais rencontraient des difficultés avec les constructions syntaxiques en anglais du fait de leur méconnaissance des patrons lexicaux. Ce constat l'a conduit à prêter plus d'attention à l'usage et à l'encodage de la langue. Hanks (2008) a résumé les principes de Hornby comme suit :

Language in use is highly patterned. Each word is typically associated with only a small number of syntactic patrons [...] Verb patterns are the most important. (Hanks, 2008, p. 2).

Cette compréhension intuitive de la nature lexicalisée du langage a fait apparaître en 1942 le premier dictionnaire d'apprentissage d'encodage destiné aux apprenants d'anglais langue étrangère, intitulé *Idiomatic and Syntactic English*<sup>4</sup> (ISED). Il proposait un inventaire systématique des patrons syntaxiques verbaux (par exemple : *Subject+Verb+Object* est un patron lié à un sens du verbe *find*), avec un codage permettant leur repérage, leur classement et leur utilisation. Pourtant, comme le souligne Hanks (2008), l'ISED manquait d'exemples authentiques nécessaires pour une étude empirique détaillée portant sur les patrons associés aux différents sens de chaque mot.

La conception de la phraséologie comme principe de fonctionnement de la langue constitue le noyau de la notion de patron essentiel pour le principe de l'idiomaticité (the idiom principle) énoncé par J. Sinclair :

The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to

---

<sup>4</sup> Ce dictionnaire comporte d'abord des entrées de verbes. Dans des éditions ultérieures les noms et les adjectifs sont aussi intégrés. Il a donné naissance au fameux dictionnaire *Oxford Advanced Learner's Dictionary of Current English* (OALD).

be analysable into segments. (Sinclair, 1991, p. 110).

Sinclair indique par ailleurs que seul le corpus permet de révéler les patrons propres au discours et propose donc une méthodologie « corpus-driven » – l’analyse empirique du corpus pour l’étude linguistique – qui constitue la méthode d’analyse fondamentale des travaux linguistiques contextualistes. Et depuis, « la lexicographie, l’enseignement et les corpus sont intimement liés. La tendance contextualiste est le fruit de l’interaction entre les trois éléments de base » (Williams, 2006, p.152).

Dans la continuité des analyses phraséologiques de Sinclair, le travail de Hunston et Francis (2000), basé sur l’analyse du corpus, consiste à identifier dans un texte les patrons distributionnels qui intègrent certaines classes de mots sémantiquement homogènes. Dans une perspective didactique, les auteurs (*ibid.*, p. 270) ont expliqué le rôle des patrons que nous pouvons résumer comme suit :

- (1) Un patron détermine le sens du mot. La plupart des mots isolés n’ont pas de sens ou plutôt des sens ambigus alors que le sens n’apparaît que dans un patron particulier qui est essentiel pour déterminer le sens du mot.
- (2) Un patron peut contribuer à la maîtrise de la langue (exactitude et variété). Suivant le principe de l’idiomaticité, le patron permet de faciliter tant la compréhension pour comprendre le sens du mot que la production pour reconstruire le sens du mot.

L’approche « The Pattern Grammar » introduite par ces auteurs a été ensuite mise en pratique dans deux volumes intitulés respectivement *Collins COBUILD Grammar Patterns 1 Verbs* et *Collins COBUILD Grammar Patterns 2 Nouns and Adjectives*, où les patrons sont « représentés presque au même niveau de généralisation, exclusivement en termes de catégories grammaticales des mots, à l’exception de certaines prépositions » (Hanks, 2008, p. 106, notre traduction). En ce qui concerne le premier volume portant sur les patrons des verbes, on remarque que soixante patrons principaux de verbes en anglais ont été dégagés. Ces patrons verbaux ont été classés sous cinq grandes catégories (Hunston & Francis, 2000, p. 51-56) :

- (1) le verbe est suivi par un groupe substantif, adjectif ou une proposition ; ex. :

*V-ing* : *She started walking*<sup>5</sup>.

- (2) le verbe est suivi par un groupe prépositionnel ou adverbial ; ex. : *V prep/adv* :  
*He ran across the road.*
- (3) le verbe est suivi par un groupe substantif et un autre élément (un autre groupe substantif, adjectif, une préposition de *that*, de *wh* ou de *-ing*) ; ex. : *V n n* : *I wrote him a letter.*
- (4) le verbe est suivi par un groupe substantif et un groupe prépositionnel ou adverbial ; ex. : *V n prep/adv* : *Stir the sugar in the bottom of the glass.*
- (5) le verbe comprend le mot *it* ; ex. : *It V clause*: *It doesn't matter what you think.*

De même que l'approche des Grammaires de Construction (celle développée par Goldberg), l'approche *Pattern Grammar* trouve son origine dans une conception phraséologique de la langue, en particulier celle portant sur l'interaction entre le lexique et la syntaxe. Nous nous référons ici aux travaux de Legallois (2005, 2006) qui ont comparé ces deux approches en étudiant la structure à deux objets directs (ex. : *give me a book*). De la manière générale, l'observation de la caractérisation sémantique des constructions et des exemples donnés dans ces deux approches montre que les significations identifiées des patrons sont plus précises et spécifiques, en d'autres termes, la constitution de classes de verbes à partir de corpus est plus fiable que celles « acquises à partir de l'expérience, générées par des archétypes conceptuels » (Legallois, 2006, p. 37). Ainsi, l'approche des Grammaires de Construction n'est pas adaptée à nos besoins et nous choisissons l'approche des patrons.

Toutefois, l'approche *Pattern Grammar* ne propose pas une analyse systématique visant à relier le sens du mot à l'usage. Dans la section suivante, nous souhaitons présenter le modèle de Hanks inscrit dans la théorie contextualiste et destiné à l'analyse lexicale en contexte. Il se distingue des approches précédentes par le fait qu'il relie le sens du mot à l'usage.

---

<sup>5</sup> Ces exemples sont repris des auteurs dans le même ouvrage.

### 1.1.3.2. Les patrons lexico-syntaxiques dans le modèle *Corpus Pattern Analysis*

Toujours dans une perspective lexicographique et didactique, Hanks (2013) propose le *Corpus Pattern Analysis* (CPA, analyse des patrons basée sur les corpus) qui est à la fois une approche sur les patrons et un modèle lexicographique basé sur corpus qui s'inscrit dans la continuité de la linguistique de corpus de Sinclair. Il s'agit d'une approche « corpus-driven » et la construction des patrons s'appuie sur l'analyse des échantillons de concordances de corpus. Hanks a pour objectif de dégager les patrons d'emploi usuel des verbes en contexte. À l'instar de Gross et de Willems, il considère le verbe comme le point d'accès à une phrase. Il souligne que le verbe détermine non seulement comment ses propres sens potentiels<sup>6</sup> sont activés, mais aussi ceux des autres éléments de la proposition. Il est donc important d'étudier comment le sens potentiel du verbe se projette sur le patron syntaxique auquel il est associé. Ici, le patron est considéré comme une configuration syntaxique intégrant un ensemble de collocations<sup>7</sup> préférées :

A pattern is a statement of the clause structure (valency) associated with a meaning of a verb—together with the typical semantic values of each argument. Different semantic values of arguments (subject, object, prepositional object) activate different meanings of the verb. (Hanks 2013 : 92).

La première application du modèle lexicographique CPA consacré à l'anglais est le projet *Pattern Dictionary of English Verbs (PDEV)*<sup>8</sup>. Ce dictionnaire, basé sur le *British National Corpus* (BNC), propose un inventaire des patrons intégrant les usages standard des verbes.

Les patrons CPA sont aussi des patrons de sens puisque chaque patron est associé à un sens (présenté sous le nom *implicature*). Afin de saisir le sens de chaque patron, les éléments lexicaux qui constituent le patron sont caractérisés sur le plan

<sup>6</sup> Selon Hanks, le mot ne possède pas de sens mais plutôt des sens potentiels qui peuvent être activés par le patron autour du mot selon des contextes différents.

<sup>7</sup> On peut se référer au sens donné par Tutin (2014, p. 30) « expressions récurrentes, essentiellement binaires, dont les éléments entretiennent une relation syntaxique comme *faire une hypothèse* ».

<sup>8</sup> Consulté le 25 novembre 2016 à l'adresse : <http://www.pdev.org.uk/#browse?q=:f=C>. Ce dictionnaire est encore en cours de construction, 1286 verbes en anglais y sont répertoriés et l'objectif est d'atteindre 5396 verbes.

sémantique à l'aide des catégories sémantiques appelées *types sémantiques*<sup>9</sup>. Ainsi, les unités lexicales intervenant en tant qu'arguments sont généralisées dans des types sémantiques qui représentent des concepts cognitifs tels [[Human]], [[Event]], [[Location]]. Ces types sémantiques sont regroupés dans une ontologie<sup>10</sup> de 200 mots établie pour le besoin de l'analyse sémantique des concordances.

L'exemple suivant illustre un patron du verbe *execute* qui signifie qu'un être humain ou une institution exécute un plan, un ordre ou une activité. [[Human]], [[Institution]], [[Plan]], [[Command]], [[Activity]] sont des types sémantiques.

Pattern<sup>11</sup> : [[Human | Institution]] execute [[Plan | Command | Activity]]<sup>12</sup> (taux de fréquence=19,2%)

Implicature : [[Human | Institution]] does work in order to put [[Plan | Command | Activity]] into effect.

Exemple : **The student** has the opportunity to formulate and execute **a search**. (Traduction de l'auteur : L'étudiant a la possibilité de concevoir et de réaliser une recherche.)

Autrement dit, dans ce patron, le verbe *execute* sélectionne de préférence des sujets noms renvoyant à un humain / une institution et des objets noms renvoyant à un plan / un ordre / une activité.

Les types sémantiques peuvent parfois être complétés par les rôles sémantiques (ex. : [[Human = Doctor]]), selon le contexte spécifique du texte. Le rôle sémantique est « le rôle que remplit un des éléments du patron en contexte » (EL Maarouf, 2011, p. 71). Considérons le patron suivant pour le verbe *swallow* (dans le sens de 'dépenser') :

Pattern : [[Activity | Institution]] swallow [[Resource = Money]] (up)

Implicature : [[Activity | Institution]] completely uses up [[Resource = Money]]

<sup>9</sup> Un type sémantique est une classe qui est attribuée à une lexie, il est indiqué entre crochets.

<sup>10</sup> Cette ontologie évolue au fur et à mesure de la construction des patrons, accessible en ligne : <http://www.pdev.org.uk/#onto>, consulté le 25 novembre 2016.

<sup>11</sup> Obtenu du site de PDEV : <http://pdev.org.uk/#browse?q=execute;f=A;v=execute>, consulté le 25 novembre 2016, un taux de fréquence est associé à chaque patron. Dans l'exemple cité, le taux de fréquence est de 19,2%.

<sup>12</sup> Lexical set [[Plan]] = {*design, plan...*} ; Lexical set [[Command]] = {*instruction, order, decision...*} ; Lexical set [[Activity]] = {*research, work, study...*} ;

Dans cet exemple, `[[resource]]` est un type sémantique alors que `[[money]]` est un rôle sémantique. Par ailleurs, certaines catégories grammaticales comme les adverbes, les prépositions ou les syntagmes grammaticaux (par exemple, les infinitifs ou les complétives) sont décrits dans les patrons lorsque ces éléments exercent une influence sur le sens du patron étudié. Par exemple, dans le patron suivant, le verbe *fail* (dans le sens de ‘ne pas parvenir à faire quelque chose’) est suivi par un syntagme prépositionnel.

Pattern : `[[Human | Institution]] fails [To-Infinitive]`

Implicature : `[[Human | Institution]] does not do something that they intended to do or were expected to do`

Le patron permet ainsi de faire le lien entre sens mobilisé et constructions verbales.

En dehors de l’anglais langue générale, l’approche CPA a aussi été adaptée à des langues romanes telles que l’espagnol dans les deux projets en cours *Diccionario de Aprendizaje de Español como Lengua Extranjera (DAELE)* (Renau & Battaner, 2012; Renau & Alonso, à paraître), *Verbario* (Nazar & Renau, 2016) et l’italien (Jezek & Frontini, 2010). De plus, il a été montré que le modèle CPA peut aussi être appliqué dans les domaines de spécialité comme l’environnement (Alonso & Renau, 2013), la médecine et la biologie (Williams & Millon, 2013).

Si l’on compare le modèle CPA au *Pattern Grammar*, la notion de patron n’est pas conçue de la même manière. On peut distinguer dans ces deux modèles les patrons lexico-syntaxiques (CPA) des patrons grammaticaux (*Pattern Grammar*) dans la mesure où les premiers incluent l’analyse sémantique des collocations intégrées dans le patron. Comme le souligne Hanks, Hunston et Francis adoptent un point de vue assez traditionnel sur la notion de patron, cohérent avec celui de Hornby, selon eux :

Patterns are composed of word classes and function words such as prepositions, and there is no role in it for statistically significant collocations. As a result, pattern grammar cannot be used to distinguish one sense of a word from another when the overall syntactic structure is the same. (Hanks, 2013, p. 6)

Nous constatons que dans le CPA, Hanks met en avant les valeurs sémantiques qui permettent de distinguer le sens lorsque deux phrases partagent la même structure syntagmatique. Comparons les deux phrases : *firing bullet from a gun* et *firing a person from a job* (Hanks, 2008, p. 21) : la *Pattern Grammar* traite ces deux phrases dans un même patron tel *V N from N* tandis que dans le CPA, deux patrons sont attribués et ils se différencient par leurs propres types sémantiques : *Fire [[Projectile]] from [[Firearm]]* vs *Fire [[Human]] from [[Activity]]*. En fait, la désambiguïsation du sens à travers les types sémantiques au niveau des compléments nous semble importante dans la mesure où elle permet non seulement de distinguer la polysémie du verbe, mais aussi de distinguer des verbes synonymes partageant la même structure syntagmatique, par exemple, *observer une évolution* (un processus), mais *constater un fait* (un résultat).

Enfin, nous constatons trois principales différences entre ces deux approches :

- (1) Le CPA est basé sur des usages réguliers des verbes et son objectif n'est pas de relever tous les patrons/sens du verbe, mais de relever les prototypes des patrons.
- (2) Le CPA examine chaque verbe individuellement et met en évidence l'association entre un patron et un sens. Ensuite, les patrons verbaux servent à distinguer les différents sens d'un verbe, tandis que le *Pattern Grammar* cherche la généralisation incluant un grand nombre de mots.
- (3) Contrairement à l'approche *Pattern Grammar* qui se limite aux catégories grammaticales, le modèle CPA considère un patron comme un ensemble incluant non seulement la structure syntagmatique, mais aussi les valeurs sémantiques. En ce sens, le CPA a accordé un rôle important aux collocations dans l'identification des patrons.

Ces caractéristiques nous permettent de justifier notre choix du modèle CPA comme approche d'analyse. En effet, les verbes étant polysémiques, y compris dans le contexte de l'écrit scientifique, l'ambiguïté peut être ainsi levée en associant les valeurs sémantiques des arguments aux structures syntaxiques. Ce modèle, accordant un rôle primordial aux contextes et aux usages en corpus, nous semble mieux adapté à notre étude avec l'objectif de faciliter la désambiguïsation dans l'encodage/décodage

du texte scientifique et d'aider la rédaction académique des apprenants non-natifs.

À la suite des approches contextualistes, nous proposons la définition de patron verbal suivante :

*Un patron verbal se compose des collocations en termes d'arguments et d'autres catégories grammaticales (préposition, adverbe, etc.) dans une structure syntaxique. Il est fréquemment associé à un sens du verbe et constitue un prototype prêt à l'usage. Seul le corpus permet de mettre en évidence les régularités lexicales et syntaxiques.*

Le patron défini dans notre étude est donc cohérent avec celui du modèle CPA. Nous nous intéressons plus spécialement aux patrons verbaux dans l'écrit scientifique. Notre domaine d'étude et notre objectif sont donc plus spécifiques. En nous inspirant du modèle CPA, nous allons adapter le modèle à notre cas et préciser quelques traitements différents dans le chapitre III. Dans ce qui suit, nous allons présenter ce que nous entendons par 'patron verbal dans l'écrit scientifique'.

## **1.2. Les écrits scientifiques et les verbes du lexique scientifique transdisciplinaire**

Dans le contexte de l'écrit scientifique, le verbe, en tant que pivot de la phrase, contribue à exprimer un grand nombre de fonctions rhétoriques et discursives telles que *formuler une hypothèse, faire la comparaison, donner une définition, ou exprimer des opinions*. Sur le plan didactique, ce lexique verbal « métascientifique et métadiscursif » (Tutin & Grossmann, 2014) est largement étudié sous l'angle transdisciplinaire dans le cadre de l'*English for Academic Purposes* (EAP) sous le nom d'« academic vocabulary » ou « sub-technical vocabulary » (Coxhead, 2002 ; Paquot, 2010). Pour le français, les études similaires dans les champs du *Français sur Objectifs Universitaires* ou *Français sur Objectifs Académiques* sont plus récentes (Cavalla, 2008, 2014) sous l'appellation de la « Langue Scientifique Générale » (LSG) (Pecman, 2005), ou du « lexique scientifique transdisciplinaire » (LST) (Tutin, 2007a, 2007b ; Drouin, 2007). Nous allons aborder le verbe du LST dans le domaine de l'écrit scientifique dans la section suivante.



### 1.2.1. Écrit scientifique et écrit académique

Un texte scientifique doit répondre aux trois critères : « une intention de connaissance explicite de l'auteur, un apport de connaissance reconnu par une communauté savante, l'inscription dans un espace de publication identifiable comme scientifique » (Berthelot, 2003, p. 33). Ces critères impliquent d'une part la scientificité qui renvoie à des pratiques intellectuelles spécifiques, et d'autre part, la reconnaissance des communautés de discours<sup>13</sup> (Bizzell, 1992) qui imposent des genres spécifiques codifiés, présentant des objectifs et des procédures communs – opérations d'analyse, de raisonnement, d'évaluation. Ces genres codifiés fondés sur le positionnement et le raisonnement scientifique existent au-delà des variations culturelles et disciplinaires, et permettent de distinguer l'écrit scientifique des écrits journalistiques et littéraires qui se caractérisent par un langage descriptif, informatif, poétique ou métaphorique (Tran, 2014). En fait, le langage formel de l'écrit scientifique se compose de traits textuels, pragmatiques, syntaxiques et lexicaux (Kocourek, 1991), ce qui pose de grandes difficultés pour les non spécialistes du domaine en question.

Partant des besoins des étudiants allophones, nous nous intéressons plus particulièrement aux écrits scientifiques en milieu universitaire (rapport, dossier, mémoire, thèse, article scientifique, communication écrite). Ces écrits relèvent donc d'un sous-domaine des écrits scientifiques, élaborés dans un contexte académique par des spécialistes d'un domaine – des experts – traitant une problématique clairement posée à laquelle ils tentent d'apporter des réponses (Swales, 1990). L'importance de la maîtrise de ces écrits est liée au fait qu'ils restent le moyen d'évaluation privilégié des études supérieures et constituent donc un défi pour les étudiants allophones en vue de valider leurs formations et obtenir un diplôme.

Nous abordons le terme « science » dans un sens large incluant les sciences exactes ainsi que les sciences humaines et sociales (SHS). Notre intérêt portant sur les

---

<sup>13</sup> La communauté de discours est définie par Bizzell (1992, p. 222) comme un groupe d'individus partageant certaines pratiques langagières. Ces pratiques sont formalisées tant au niveau des conventions stylistiques que des savoirs scientifiques.

disciplines SHS s'explique par la demande des étudiants et la spécificité/l'exigence des écrits de ces disciplines. Tout d'abord, comme le constate Omer (2014), le taux d'échec des étudiants étrangers inscrits en SHS est deux fois supérieur à celui des Français, ce qui est essentiellement imputable aux compétences langagières des étudiants (Mangiante & Parpette, 2011). En effet, les besoins linguistiques sont les plus manifestes en SHS : d'une part, les disciplines des SHS se caractérisent par un langage métadiscursif et métascientifique qui impose une structuration textuelle spécifique (Tran, 2014); d'autre part, ces disciplines accordent une grande importance à l'évaluation de la pratique du genre textuel. Cette évaluation porte tant sur la forme que sur le fond, ce qui oblige l'étudiant à faire preuve d'une grande aisance rédactionnelle au croisement de plusieurs savoirs : savoir disciplinaire (connaissances scientifiques liées à la discipline), savoir sur la structure organisationnelle, connaissance du lexique (Cavalla, 2007).

Dans notre étude, nous nous focalisons sur le lexique des écrits scientifiques. Cette étude est envisagée dans une perspective large qui implique non seulement les aspects lexicaux, mais aussi discursifs.

### **1.2.2. Le lexique scientifique transdisciplinaire (LST) et le projet**

#### **Scientext**

Dans le domaine de l'écrit scientifique, plusieurs auteurs s'accordent sur l'existence d'une langue générale scientifique, « partagée par la communauté scientifique mise en œuvre dans la description et la présentation de l'activité scientifique [...] qui ne renvoie pas aux objets scientifiques des domaines de spécialité, mais au discours sur les objets et les procédures scientifiques » (Tutin, 2007a, p. 6). Elle reçoit différentes appellations : le *Vocabulaire Général d'Orientation Scientifique* (VGOS) (Phal, 1971), *Vocabulaire général d'enseignement scientifique* (VGES) (Chetouani, 1997), la *Langue Scientifique Générale* (LSG) (Pecman, 2004). Ces études débouchent sur la construction de listes de mots de la langue générale scientifique qui sont basées essentiellement sur des corpus des disciplines en sciences exactes. Comme nous l'avons souligné plus haut, notre intérêt porte sur les disciplines SHS, c'est pourquoi nous nous inscrivons dans la lignée des

travaux de Tutin (2007a) dont l'étude se fonde sur un corpus composé d'écrits scientifiques venant des disciplines SHS dans le corpus Scientext.

Selon elle, l'écrit scientifique est composé de cinq strates de lexique : le lexique propre aux écrits scientifiques (traversant : *analyse, proposer un modèle, mener une enquête*), le lexique abstrait général (non spécifique aux écrits scientifiques et relevant aussi d'autres types d'écrits argumentatifs ou informatifs : *voir l'intérêt, autrement dit, à savoir, considérer...comme...*), le lexique méthodologique disciplinaire (métascientifique, spécialisé et lié à la discipline : *en coupe instantanée, panel* ), le lexique terminologique (*collocation, biophysique*) et le lexique de la langue générale (*connaître, habitudes de consommation*).

À la suite de Tutin, nous nous intéressons plus particulièrement aux deux premiers types, à savoir le lexique propre aux écrits scientifiques et le lexique abstrait général. Ce choix se justifie dans notre objectif didactique. Les enjeux liés aux écrits scientifiques nous conduisent à constater qu'en matière d'écrits scientifiques, la principale difficulté des étudiants étrangers réside dans ce lexique intermédiaire entre la langue générale et la terminologie. Il s'agit du lexique qui décrit les activités scientifiques : de l'exposition du cadre de la recherche aux résultats en passant par la méthodologie et l'argumentation. Ce lexique appelé « lexique scientifique transdisciplinaire » (LST) (Tutin, 2007b) constitue le fonds commun lexical propre à la démonstration et au raisonnement scientifiques dont l'apparition au sein de disciplines multiples ne doit pas masquer les différences d'emploi. À la suite de Tutin, nous avons dégagé plusieurs caractéristiques liées à ce lexique.

D'abord, c'est un lexique transversal aux différentes disciplines. Contrairement aux mots appartenant à la discipline spécialisée, le LST est transversal dans la mesure où les procédures, les objets et les concepts des activités scientifiques sont partagés par les disciplines, avec parfois des acceptions légèrement différentes (ex : *théorique, comparaison, analyser*).

Ensuite, le LST n'est pas un lexique spécialisé, mais plutôt un lexique abstrait particulièrement représenté dans les écrits scientifiques » (Tutin, 2007a) (ex : *poser un problème, permettre de*). En fait, d'une part, il est difficile de dissocier ce lexique, spécifique aux écrits scientifiques, du lexique abstrait non terminologique qui peut

relever d'autres types d'écrits argumentatifs ou informatifs comme l'écrit journalistique. D'autre part, toujours dans une perspective didactique, il est intéressant d'aborder le lexique spécifique à l'écrit scientifique, mais aussi le lexique surreprésenté dans ce type d'écrit. Dans la lignée des travaux consacrés à l'extraction du LST, la plupart des chercheurs intègrent un ensemble de mots abstraits dont un large sous-ensemble n'a rien d'exclusivement scientifique (Phal, 1970 ; Pecman, 2007 ; Paquot, 2010 ; Hatier, 2013).

Une autre caractéristique de ce lexique est qu'il est fréquent et omniprésent. Enfin, il intègre une importante phraséologie. En effet, il apparaît dans des collocations lexicales (ex : *formuler une hypothèse, vérifier les résultats, rejeter une hypothèse,*) ou des constructions syntaxiques préférentielles (ex : *on pourrait considérer que, comme nous l'avons indiqué ci-dessous, cela ne signifie pas pour autant que*).

D'après Tutin (2014, p. 27), la phraséologie par son caractère préconstruit permet à l'auteur/chercheur de « renforcer son appartenance à une communauté de discours » et facilite ainsi l'écriture. De plus, sa récurrence dans différentes disciplines est « une voie d'entrée privilégiée pour l'étude épistémologique des discours scientifiques ». Comme le souligne Wray (2002), cette phraséologie, en particulier sa fonctionnalité rhétorique ou discursive, constitue une condition d'accès aux compétences lexico-grammaticales des locuteurs et contribue ainsi à la promotion de l'intérêt personnel du locuteur « exprimer l'information avec fluidité, organiser l'information, signaler la structure du texte, aider à la planification du discours, agir sur l'interlocuteur, acquérir une information demandée, être perçu comme un membre du groupe, etc. » (Legallois & Tutin, 2013, p. 19).

Dans cette perspective, la phraséologie est au cœur de l'étude du LST dans le discours scientifique. De nombreux travaux (Boch & Rinck, 2010 ; Cavalla, 2010 ; Tutin & Grossmann, 2014) autour du projet Scientext sont consacrés à l'étude de la phraséologie dans les marqueurs lexicaux, grammaticaux et énonciatifs, sur les thèmes du positionnement et du raisonnement. Le verbe y occupe une place importante. Mais il est surtout étudié à travers les collocations de type V-N (Grossmann & Tutin, 2010a ; Cavalla, 2008 ; Chavez, 2008) et l'aspect syntaxique est

encore peu étudié sauf dans quelques études de cas : verbes de constat (*on peut constater que*) (Grossmann, 2014b), verbes causatifs (*permettre de proposer que, amener à faire une hypothèse*) (Bak & Novakova, 2013), constructions verbales impersonnelles et passives (*il ressort de ces travaux que, peut être considéré comme*) (Fifielska, 2015). Les patrons des verbes du LST dans l'écrit scientifique n'ont pas – à notre connaissance – fait l'objet de travaux en français que ce soit dans la linguistique ou la didactique, d'où l'intérêt de notre étude.

En revanche, en anglais, les verbes de la langue générale scientifique sont largement étudiés : listes et classements par des critères différents (fréquence, sémantique ou fonctions rhétoriques) ainsi que des travaux sur les patrons des groupes de verbes spécifiques.

### 1.2.3. Verbes dans l'*English for Academic Purposes*

Comme nous l'avons mentionné plus haut, le LST rejoint le concept similaire du vocabulaire anglais étudié dans le cadre du *English for Academic Purposes* (ci-après désigné par EAP). Dans la section suivante, il nous semble donc important de présenter quelques-uns des nombreux travaux effectués sur ce lexique dans le domaine EAP, concernant notamment les questions liées à la délimitation et à la classification des mots du EAP.

#### 1.2.3.1. L'*Academic word List* de Coxhead

Coxhead (2010, 2002) a défini les mots du EAP dans *Academic Word List* (ci-après désigné par AWL). Cette liste organisée par niveaux de fréquence regroupe 570 familles de mots du EAP les plus fréquents répartis dans 28 disciplines dont le droit, les sciences, la littérature, et le commerce comme *analyse, approach, concept, context, formula, method* etc. Les sens de ces mots sont considérés comme non spécifiques par rapport à la terminologie d'une discipline scientifique, mais en même temps, plus formels que la langue générale. En revanche, on peut souligner, à l'instar de Tran (2014), que l'étude de Coxhead n'est basée que sur des critères lexicométriques, des indices de fréquence et n'effectue pas encore un classement notionnel. Pour l'exploitation pédagogique de cette liste, il nous semble indispensable

d'« intégrer des informations sur les fonctions sémantiques et rhétoriques » (Tutin, 2010c, p. 145).

Ce point de vue est partagé par Paquot (2010) qui confirme qu'il faut se focaliser sur les fonctions discursives et rhétoriques de l'« academic vocabulary » en ajoutant que le champ d'étude restant ouvert, l'« academic vocabulary » ne peut pas être considéré comme une liste de mots spécialisés exclusifs.

### 1.2.3.2. L'*Academic Keyword List* de Paquot

L'*Academic Keyword List* (ci-après désigné par AKL) de Paquot est basée sur le critère de fréquence, mais aussi sur les critères sémantiques et pragmatiques. Par exemple, le nom *analysis* étant morphologiquement associé au verbe *analyse* est ajouté même s'il n'atteint pas le critère de fréquence attendu (*ibid.*, p. 54-55). Cette liste couvre 40% de l'AWL et inclut des mots qui en sont absents comme *aim*, *argue*, *cause*, *claim*, *effect*. Au total, elle regroupe 930 mots dont 233 verbes<sup>14</sup>, répartis essentiellement en sciences humaines et sociales. Il faut noter par ailleurs que les mots sont relevés dans un grand corpus de textes académiques (dont des extraits du *BNC* et dans les corpus de différentes disciplines) ainsi que dans des corpus d'apprenants dont la langue maternelle est, pour une grande majorité, le français.

La plupart de ces mots peuvent être classés en 12 fonctions rhétoriques dont illustration, cause et effet, comparaison et contraste, concession, ajout de l'information, expression de l'opinion, expression de la possibilité et de la certitude, introduction du sujet et des idées, énumération, reformulation, citation et conclusion. Concernant les verbes, ils assurent la plupart de ces fonctions, entre autres, ceux qui expriment l'illustration tels que *illustrate*, *exemplify*, etc. Ces fonctions rhétoriques nous semblent intéressantes sur le plan didactique et nous inspirent dans notre travail sur le classement des patrons de verbes. De plus, cette approche fonctionnelle est appliquée dans le dictionnaire le *Louvain English Academic Purposes Dictionary* (LEAD) (Granger & Paquot, 2010) que nous allons présenter dans la section 1.4.2.

Dans ces deux listes, les verbes figurent toujours à une place prédominante du

<sup>14</sup> Consulté le 14 janvier 2016 à l'adresse : <https://www.uclouvain.be/en-372126.html>

fait de leur co-occurrence avec les noms, les adverbes et de leur apparition récurrente dans des structures syntaxiques. Ils font l'objet de travaux de classification dans lesquels les chercheurs s'accordent à prendre en compte leurs propriétés sémantiques et syntaxiques.

### 1.2.3.3. Le classement des verbes du EAP

Dans notre étude, nous souhaitons établir un classement des verbes du LST en fonction de leurs propriétés sémantiques et syntaxiques (*cf.* Chapitre III). Faute de référence en français, nous allons notamment nous référer aux travaux similaires effectués en anglais qui peuvent inspirer notre travail de classement.

Dans une perspective d'analyse de l'écrit académique à visée didactique, Hinkel (2004) propose une classification des verbes du EAP en prenant en compte les fonctions que jouent ces verbes dans l'écrit scientifique :

- (1) Verbes d'activités (activity verbs) qui renvoient aux actions volontaires ou non volontaires comme *obtain, use, produce, give*.
- (2) Verbes psychologiques (mental/emotive verbs) qui représentent les états intellectuels ou les processus cognitifs tels *see, find, consider*.
- (3) Verbes de déclaration (reporting verbs) qui se rapportent aux opinions comme *suggest, discuss, argue, emphasize*.
- (4) Verbes de liaison (linking verbs) qui relèvent des états ou des existences comme *be, become, keep, prove*.
- (5) Verbes de relation logique-sémantique (logico-semantic relationship verbs) qui impliquent la cause/l'effet ou un changement comme *cause, involve, illustrate*.

L'auteure précise que ces cinq classes sémantiques jouent un rôle important dans la rédaction scientifique. À côté de la liste, des pistes didactiques sont proposées sur l'enseignement de ces classes de verbes, et l'auteure souligne l'importance des combinaisons lexicales (les noms, les prépositions et les complétives) et les

régularités syntaxiques.

That is, by simply replacing the subject noun or the verb with various proximate alternatives, L2 writers can actually come up with a great number of diverse and contextually appropriate expressions. (Hinkel, 2004, p. 203)

Par exemple, l’auteure souligne que les verbes de déclaration ont tendance à apparaître dans certaines constructions telles que *The author(s) V that, The author(s) describe N (data, information, etc.) to V that*. Selon elle, ces régularités syntaxiques doivent être mises en avant pour aider l’apprenant à s’approprier des expressions adaptées dans leur écrit académique. Nous partageons l’avis de l’auteur sur cette réflexion. En revanche, il nous semble que cette classification peut être affinée et subdivisée en différents champs sémantiques pour faire ressortir des propriétés sémantiques et syntaxiques plus homogènes.

Dans le cadre du projet KIAP<sup>15</sup>, Fløttum et ses collègues (Fløttum et *al.*, 2006) en se basant sur un corpus varié d’écrits scientifiques en sciences humaines (linguistique), sciences sociales (économie) et sciences expérimentales (médecine), ont étudié les valeurs sémantiques et pragmatiques des verbes afin de déterminer les rôles et les activités de l’auteur dans l’écrit scientifique. Quatre groupes de verbes correspondant à quatre rôles (scripteur, chercheur, argumentateur, évaluateur) sont proposés et basés sur les fonctions discursives suivantes :

- (1) Verbes de discours (discourse verbs) liés aux activités de la rédaction scientifique : *present, show, discuss, etc.* Cette classe peut correspondre à la classe 1 (verbes d’activités) de Hinkel.
- (2) Verbes de recherche (research verbs) liés aux activités de la recherche : *analyse, compare, explore, use, etc.* Cette classe peut correspondre à la classe 1 (verbes d’activités) de Hinkel.
- (3) Verbes de position (position verbs) liés aux activités de l’argumentation : *argue, claim, propose, etc.* Cette classe est proche de la classe 3 (verbes de déclaration) de Hinkel.
- (4) Verbes/constructions verbales d’évaluation et d’émotion (evaluation and

<sup>15</sup> KIAP : Kulturell Identitet i Akademisk Prosa (Identité culturelle dans les écrits scientifiques).



emotion verbs and verb constructions) liés aux activités de l'évaluation : *feel*, *be struck by*, *be content to*, etc. Cette classe est proche de la classe 2 (verbes psychologiques) de Hinkel.

D'une manière générale, par rapport à la classification proposée par Hinkel, Fløttum n'a pas inclus les classes des verbes ayant un sujet inanimé (verbes de liaison et verbes de relation logique-sémantique), car elle se limite aux structures avec le pronom de la première personne. En outre, dans la classification de Fløttum, les propriétés sémantiques et syntaxiques des verbes ne jouent qu'un rôle secondaire, ce sont les activités de l'auteur qui priment sur le sens du verbe, par exemple, dans les deux phrases suivantes :

- A. In section 4 below, **I show how** force-dynamic relations constitute a sufficient condition for dynamicity...
- B. **I show that** two recent accounts of this are inadequate, both conceptually... (Fløttum et al., 2006, p. 89)

D'après Fløttum, le premier verbe *show* appartient aux verbes de discours alors que le deuxième appartient aux verbes de position. Ceci s'explique par le fait que, dans la première phrase, le verbe apparaît dans un co-texte métatextuel (*in section 4 below*) qui présente la structure du texte. Cette distinction ne nous semble pas utile puisque nous envisageons de proposer un classement des verbes en fonction de leurs propriétés sémantiques et syntaxiques.

Enfin, contrairement à Hinkel, Fløttum a abordé la question de la polysémie des verbes du EAP dans sa classification. Elle souligne que certains verbes peuvent entrer dans plusieurs groupes. Par exemple, le verbe *find* dans l'exemple (*We've found three types of verbs.*) appartient aux verbes de recherche qui servent à décrire les activités scientifiques alors que dans l'exemple (*I have found this problem very complicated.*), le verbe prend le sens d'un verbe d'évaluation qui porte sur les objets scientifiques. Dans notre étude, nous postulons que l'analyse des propriétés sémantiques et syntaxiques des verbes du LST permettrait de prendre en compte la question de la polysémie.

#### 1.2.3.4. La phraséologie des verbes du EAP

En dehors des travaux sur le classement des verbes du EAP, ces verbes sont surtout étudiés dans une notion étendue de la phraséologie héritière de la linguistique de corpus anglo-saxonne avec des approches statistiques. Selon Bolly (2011), les unités polylexicales relevées par les études de corpus en phraséologie statistique se présentent sous différentes formes :

Une combinaison de deux ou plusieurs unités linguistiques liées entre elles lexicalement (par exemple, les co-occurrences et les séquences récurrentes), sémantiquement (par exemple, la prosodie sémantique) et/ou syntaxiquement (par exemple, les colligations, les patrons, les cadres collocationnels et les collostructions). Elles se caractérisent par une récurrence significative de co-apparition dans l'usage. (Bolly, 2011, p. 54)

Ce propos nous amène à supposer que le concept clé qui relie ces différentes unités polylexicales réside dans la relation de co-apparition des mots dans le contexte, celle-ci peut être interprétée comme suit :

the relationship that a lexical item has with items that appear with greater than random probability in its textual context. (Hoey, 1991, p. 6)

Même si chaque discipline peut montrer sa propre préférence pour certaines unités polylexicales particulières (Granger & Paquot, 2009a), les chercheurs s'intéressent davantage à celles qui assurent des fonctions métalinguistiques et métadiscursifs, communs à toutes les disciplines : *as can be seen* (Hyland, 2008a), *has been frequently described as* (Gledhill, 2011). Ces unités polylexicales sont souvent associées aux fonctions rhétoriques, comme le souligne l'étude de Hyland (2008a) : présenter la structure (*as shown in figure*), indiquer les résultats (*these results suggest that*), guider le lecteur dans la lecture (*it should be noted that, as can be seen*). S'il est admis que les verbes du EAP ont une préférence d'apparition dans certaines associations lexico-grammaticales, nous souhaitons examiner dans la section suivante les travaux sur les patrons lexico-syntaxiques des verbes du EAP/LST et illustrer les patrons qui nous intéressent.

### 1.2.4. Les patrons des verbes du LST dans le cadre du CPA

Ayant comme objectif d'observer le transfert des unités lexicales de la langue générale à la langue de spécialité et de proposer un dictionnaire spécialisé aux apprenants de l'anglais langue étrangère, Williams *et al.* (2011a) appliquent la méthodologie du CPA à un corpus spécialisé (essentiellement les disciplines biologie et médecine) de 33 millions de mots - *BioMed Central corpus* (BMC) pour établir les patrons des verbes les plus fréquents en les comparant à ceux du PDEV de Hanks, comme dans l'exemple du verbe *treat* ci-dessous :

BMC : [[Human 1 | Human Group]] treat [[Human 2 = Patient | Laboratory Animal = Rat, Mouse | Organism= Cell]] (with [[Drug= Vehicle]])

PDEV : [[Human 1 | Process | Drug]] treats [[Human 2 | Animal | Disease | Injury]]

Comme le soulignent les auteurs (Williams, Millon, & Alonso, 2011, p. 18), c'est surtout la sélection des types/rôles sémantiques ainsi que des groupes lexicaux spécifiques au domaine d'étude qui marque la différence entre le sens général et le sens spécifique du verbe.

By looking at the concordances, it can be observed that even though the patterns are similar to that used in general language, in the BMC the difference lying in the arguments of the verb; in terms of Hanks, on the number of collocates that normally populate an argument slot in relation to the verb. While in general language '[[Anything]] is called', in our corpus 'X' is a more limited set of lexical items. (Williams, Millon, & Alonso, 2012, p. 1011- 1012)

Par exemple, dans le BMC, le type sémantique [[Laboratory Animal = Rat, Mouse]] est plus spécifique que [[Animal]]. Cette sélection particulière conduit à un changement de sens du général au spécifique. Dans l'exemple du verbe *treat*, un animal n'est pas guéri, mais il fait l'objet d'une recherche médicale qui vise à obtenir un remède à une maladie. Cette étude nous semble intéressante dans la mesure où elle montre que le modèle CPA permet de mettre en évidence les usages prototypiques des verbes dans le domaine scientifique. En revanche, il faut noter que cette étude ne s'intéresse pas qu'aux verbes du LST, mais à tous les verbes présents dans l'écrit scientifique. Nous présentons plus en détail le dictionnaire spécialisé conçu par

Williams et ses collègues dans la section 1.4.3.

Schutz (2012) applique le modèle CPA pour décrire des verbes polysémiques comme *find* dans l'écrit académique de l'anglais. Son corpus de 3 millions de mots rassemble trois disciplines, le commerce, la linguistique et la médecine. Les patrons lexico-syntaxiques du verbe *find* dans ces trois disciplines ont été identifiés et le résultat montre que les patrons les plus fréquents sont communs aux trois disciplines :

- (1) [[HUMAN|STUDY]] find [[{that}-clause]]
  - Carpenter found that** stakeholder characteristics influenced global risk taking of IPO firms. (commerce)
  - Corpus linguists** have **found that** affective meaning can extend over segments of text larger than an orthographic word. (linguistique)
  - Some studies found that** those who drank coffee were the healthiest cohort members...(médecine)
  
- (2) [[HUMAN|STUDY]] find [[RELATIONSHIP]] {between} [[ANYTHING]]
  - Bellu found a correlation between** task motivation and sales growth. (commerce)
  - Studies found a relationship between** frequency of exposure and acquisition, in particular for beginning learners. (linguistique)
  - Studies** have consistently **found an association between** higher coffee consumption and lower risk for type 2 diabetes. (médecine)

L'auteure a par ailleurs noté que, pour les patrons les moins fréquents, la grande différence entre les disciplines se situe au niveau de collocations, d'où l'intérêt de l'enseignement de ces collocations relevant de la discipline :

- (1) Be found in [STUDY|DATA|HUMAN|GROUP...]
  - Process networks **are** typically **found in large organizations**. (commerce)
  - This use of *wis* **was** not **found in our corpus**. (linguistique)
  - These effects **were** not **found in patients with meningitis**. (médecine)

Cette étude nous semble particulièrement intéressante parce qu'elle consiste à appliquer le CPA dans le domaine scientifique d'un point de vue transdisciplinaire, ce qui est proche de notre objectif. Cependant, il faut noter que la discipline de la médecine, qui appartient au domaine des « sciences expérimentales », est éloignée

des deux autres. Nous supposons que l'étude des patrons de verbes au sein des disciplines SHS apporte des résultats plus homogènes, ce qui nécessite un travail de généralisation rigoureux.

Rappelons que nous étudions la phraséologie dans une conception large qui porte sur les patrons de verbes dans l'usage, lexicaux et lexico-syntaxiques, repérables par le biais d'analyses de fréquence en linguistique de corpus. Nous allons illustrer les verbes du LST ainsi que les patrons des verbes qui nous intéressent, à partir d'un exemple :

Au cours de cette **contribution**, nous avons **tenté d'établir** un lien entre la **théorie** économique du consommateur et le **concept** de développement soutenable. **Pour ce faire**, nous avons **procédé en** deux étapes. **Dans un premier temps**, nous **proposons un aperçu** de la consommation dans la **théorie** économique, en partant de l'**analyse** économique standard pour mettre en avant ses **limites...**  
[267-Article-Économie-Conclusion]

**Figure 1-1 : Les verbes du LST et les patrons verbaux dans une conclusion d'article d'économie**

Les mots du LST sont marqués en gras, les mots constituant des patrons des verbes sont soulignés. Nous présentons ci-dessous les 3 patrons potentiels.

- (1) nous - tenter de - établir lien
- (2) nous - procéder - en étapes
- (3) nous - proposer - un aperçu

Bien évidemment, la construction d'un patron lexico-syntaxique est le résultat d'une modélisation linguistique qui demande de prendre en considération toutes les lignes des concordances du mot dans le corpus. Nous présenterons notre méthode de la modélisation des patrons dans le chapitre III.

Comme on peut le constater, les verbes du LST assurent des fonctions essentielles dans l'argumentation, l'énonciation, et l'organisation du discours. Ils ont tendance à apparaître dans des patrons lexico-syntaxiques. En fait, dans leurs écrits scientifiques, les apprenants doivent accorder une grande importance à ce contexte

phraséologique relevant de l'organisation du discours et des activités scientifiques qui leur posent de grandes difficultés. Afin de justifier notre propos, nous examinerons dans le chapitre V l'utilisation de ces éléments phraséologiques dans un corpus d'apprenants.

### **1.3. La phraséologie du lexique scientifique transdisciplinaire : déficits persistants des apprenants non-natifs**

Nous avons déjà montré que la phraséologie dont font partie les patrons verbaux occupe une place de premier plan dans l'écrit scientifique. Comme le souligne Tutin (2014), d'une part, la maîtrise de ces éléments « prêts à écrire » permet aux apprenants d'entrer dans la « communauté de discours » ; d'autre part, la phraséologie transdisciplinaire n'est pas une terminologie, mais peut être considérée comme un lexique du genre. De ce fait, elle « constitue une voie d'entrée privilégiée pour l'étude épistémologique des discours scientifiques » (*ibid.*, p. 27).

En revanche, dans le cadre didactique, de nombreux travaux menés dans le domaine de l'*English as a Foreign Language* (ci-après désigné par EFL) (Nesselhauf, 2005 ; Granger & Paquot, 2009b) ont montré que les apprenants d'une langue étrangère peinent à manier les éléments phraséologiques (allant des collocations aux routines phraséologiques) d'une manière fluide et experte. Il en va de même pour le domaine du français langue étrangère (ci-après désigné par FLE) (Cavalla, 2009 ; 2015 ; Bolly, 2011), mais faute de réels corpus d'apprenants du FLE en France, on constate que les recherches restent beaucoup moins nombreuses (Forsberg, 2006).

En fait, la plupart des travaux sur l'utilisation des phénomènes phraséologiques chez les apprenants sont basés sur les données collectées auprès d'apprenants, à l'oral comme à l'écrit. On voit donc l'apparition de différents corpus d'apprenants (Granger *et al.*, 2015) (*cf.* Chapitre II) qui permettent d'une part de relever les erreurs dans les productions des apprenants, et d'autre part d'« informer les outils de référence pour cibler des éléments de langue sous- ou surexploités (comparés aux productions de natifs ou d'autres apprenants) » (Boulton *et al.*, 2013, p. 24).

Comme souligné plus haut, les travaux didactiques ont été menés sur différents

types d'éléments phraséologiques, en particulier les collocations (Nesselhauf, 2005) et les séquences récurrentes (Lake, 2004 ; Paquot, 2010). Il faut souligner que ces éléments phraséologiques restent la source d'erreurs jusqu'à un niveau avancé d'apprentissage, comme le montrent deux études intéressantes dans le domaine du FLE. Cavalla (2015) a relevé des exemples de collocations erronées du type V+N rencontrées dans les écrits des doctorants non-natifs du FLE : *mettre l'hypothèse* (*émettre une hypothèse*), *continuer à investiguer cette piste de recherche* (*poursuivre une investigation*). Par ailleurs, elle a remarqué que certains doctorants ont un répertoire limité de collocations et qu'ils utilisent des périphrases pour exprimer une idée. Par exemple, au lieu d'employer la collocation *suivre des cours/heures de X*, un étudiant a écrit *faire plus qu'une centaine d'heures de français*.

Toujours dans le domaine des écrits académiques, Bolly (2011) a étudié 589 séquences récurrentes construites avec les deux verbes à haute fréquence *prendre* et *donner* dans les textes académiques des apprenants du français L2 anglophones et néerlandophones. L'auteur a montré que les apprenants non-natifs produisent essentiellement des « déviations formelles » qui sont liées au caractère non acceptable ou non attesté des unités phraséologiques dans l'usage. Il nous semble intéressant de noter que le verbe est la catégorie grammaticale qui présente le plus d'erreurs - 25% des unités phraséologiques déviantes : *prendre une vue microéconomique* (*adopter un point de vue microéconomique*), *la réflexion que donne le théâtre* (*la réflexion qu'apporte le théâtre*).

En dehors des erreurs, les chercheurs dans le domaine de l'EFL ont également relevé des cas de suremploi et de sous-emploi chez les apprenants par rapport aux natifs, par exemple, le suremploi de certaines séquences peu fréquentes chez les natifs telles que *as far as I am concerned* chez les apprenants français (Granger, 2008) , *to my mind, I think that* chez les apprenants de différentes langues maternelles (Gilquin & Paquot, 2008). Les cas de sous-emploi s'avèrent plus nombreux : des séquences remplissant des fonctions rhétoriques importantes comme *it's possible that* chez les apprenants sinophones (HongKong) (Hyland, 2008b), *as suggested by* (Granger & Paquot, 2009b) et *it could be argued that, be explained in terms of* (Paquot, 2010) chez les apprenants de différentes langues maternelles.

Parmi les éléments phraséologiques, l'utilisation des patrons verbaux constitue un des déficits les plus persistants, du fait que ces unités sont linguistiquement complexes (Bolly, 2011) : elles relèvent à la fois de la syntaxe et du lexique – préférences sémantiques, alternances syntaxiques, la polysémie des verbes. Par ailleurs, on s'aperçoit qu'il y a très peu d'études de ce type en FLE.

Dans une étude précédente (Hatier & Yan, 2017), nous avons examiné les constructions verbales de 4 verbes du LST (*montrer, considérer, décrire et expliquer*) dans un corpus composé de mémoires de recherche d'un groupe d'apprenants chinois<sup>16</sup> en le comparant au corpus Scientext-SHS<sup>17</sup>. Les résultats montrent que par rapport aux experts, les constructions verbales présentes dans le corpus d'apprenants se caractérisent par une sous-utilisation :

1. de certaines alternances syntaxiques : passif, passif pronominal ;
2. d'éléments lexicaux : noms du LST, pronom sujet *on* et *nous* ;
3. de certaines constructions spécifiques, en particulier celles qui servent à aborder un sujet (*si l'on considère*), à guider le lecteur dans le plan du texte (*comme nous l'avons montré plus haut*) et à établir les liens entre les éléments textuels (*ceci s'explique par*).

À la suite de cette étude, nous allons approfondir l'analyse des difficultés des apprenants chinois en matière de constructions verbales dans le chapitre V où l'on propose un diagnostic tant sur les erreurs que sur les cas de suremploi et de sous-emploi.

#### **1.4. Problématisation et réflexion en vue des applications pédagogiques sur les patrons verbaux**

Les études sur les phénomènes phraséologiques peuvent conduire à des applications pédagogiques concrètes. Comme le souligne Williams (2008a), le besoin

<sup>16</sup> Nous envisageons d'intégrer à cette étude un corpus d'écrits d'étudiants natifs (corpus de littéracie avancée) afin de le comparer aux corpus d'experts et de non-natifs.

<sup>17</sup> Composé d'articles de recherche de revue en sciences humaines et sociales.



d'une ressource spécialisée, rendant compte des unités préfabriquées (collocations, patrons) pour faciliter la rédaction scientifique des apprenants, est bien réel. Dans cette partie, nous allons étudier quelques ressources centrées sur les éléments phraséologiques : le dictionnaire multilingue *Aide à la Rédaction de TExtes Scientifiques* (ARTES) (Pecman & Kübler, 2011), le *Louvain English Academic Purposes Dictionary* (LEAD) (Granger & Paquot, 2010) et l'*E-Advanced Learner's Dictionary of Verbs in Science* (DicSci) (Williams *et al.*, 2012). Notons par ailleurs qu'à notre connaissance, il n'existe pas encore d'outil spécifique pour aider les apprenants du FLE dans leur rédaction académique/scientifique.

L'étude des ressources pédagogiques va guider notre analyse linguistique ainsi que nos réflexions didactiques.

#### **1.4.1. Combiner l'approche syntactico-sémantique et onomasiologique : le modèle de Pecman et son application dans l'*Aide à la Rédaction de TExtes Scientifiques* (ARTES)**

Le dictionnaire ARTES<sup>18</sup> est une base de données terminologique et phraséologique propre à la langue scientifique (Pecman & Kübler, 2011), destinée à faciliter la rédaction scientifique et la traduction dans les discours spécialisés. La base en ligne propose notamment deux fonctions de recherche des données : la première d'ordre terminologique (la rubrique « terminologie en contexte ») permet d'accéder aux termes spécialisés avec les informations sur les définitions, les collocations et les traductions ; la deuxième d'ordre phraséologique (la rubrique « phraséologie discursive »), basée principalement sur l'étude de Pecman (2004, 2007, 2012), fournit les unités phraséologiques<sup>19</sup> (notamment en anglais et en français) de la *Langue*

<sup>18</sup> La base de données a été conçue dans une perspective plurilingue (français, anglais, allemand, espagnol, etc.) et a accordé une priorité au français et à l'anglais. La ressource lexicale concerne notamment les domaines des Sciences de la Terre et des Planètes (STEP) et les domaines des sciences de l'environnement. Elle est en grande partie en libre accès : <https://artes.eila.univ-paris-diderot.fr/>, consulté le 15 mai 2016.

<sup>19</sup> D'après Pecman (2004), les unités phraséologiques peuvent être définies comme « toute association arbitraire de deux ou plusieurs mots susceptible d'être réutilisée à droit au statut d'unité phraséologique » (Pecman, 2004, p. 28), par exemple, *jouer un rôle, objectif ultime, il est possible de,*

---

*Scientifique Générale* (LSG) organisées selon différentes fonctions rhétoriques/discursives<sup>20</sup>.

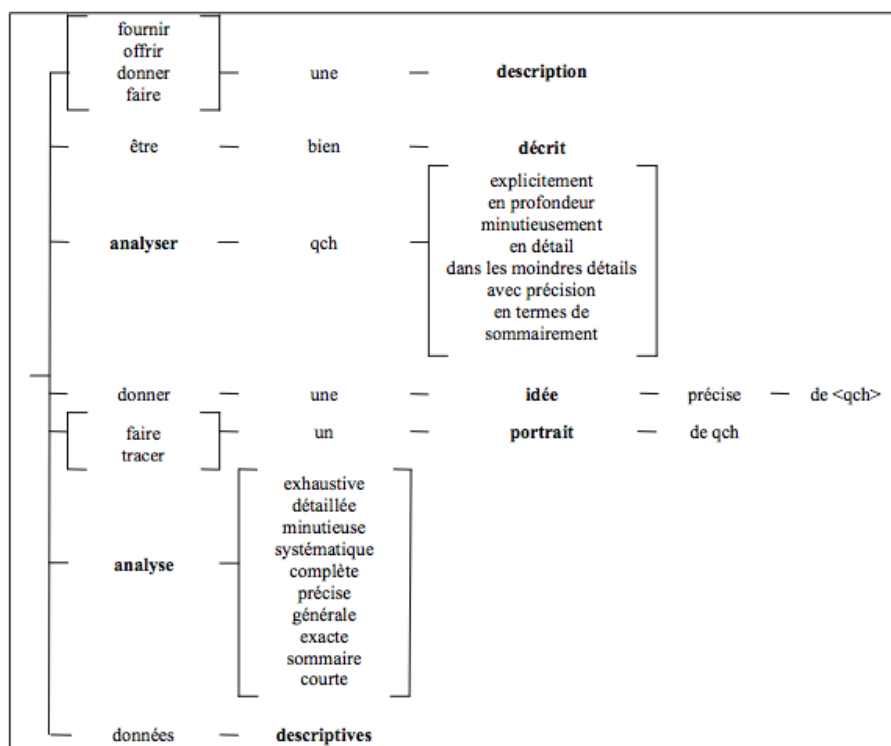
Ici, nous nous intéressons plus particulièrement à l'entrée par la phraséologie discursive. Dans le cadre du projet ARTES, à partir d'un corpus aligné (français-anglais), composé des textes scientifiques des disciplines sus-citées, Pecman (2012) a examiné les spécificités des unités phraséologiques correspondant aux différentes fonctions rhétoriques et discursives codées des textes spécialisés. D'après elle, ces éléments phraséologiques jouent un rôle important dans la construction d'un discours et d'un genre. Elle a souligné par ailleurs que ces fonctions propres au genre textuel sont notamment non spécifiques aux domaines.

En combinant l'approche syntactico-sémantique et l'approche onomasiologique, Pecman (2007) a identifié des notions clés de la LSG telles que |DESCRIPTION|, |CONSEQUENCE|, |OBSERVATION|, etc. (au total 125 notions relevées). Chaque notion est reliée à plusieurs unités phraséologiques sémantiquement apparentées. Les différentes structures collocationnelles relevant de chaque notion sont analysées afin de construire le schéma combinatoire. Par exemple, la figure 1-2 illustre un extrait du schéma combinatoire des unités phraséologiques ayant trait à la notion |DESCRIPTION|.

---

etc.

<sup>20</sup> Les fonctions rhétoriques/discursives sont définies par Pecman (2012) comme des « unités textuelles au contenu informationnel attendu, qui s'actualisent dans des structures lexico-grammaticales reconnaissables au sein d'un type de discours, et, nécessairement, d'un type de genre textuel » (Pecman, 2012, p. 118).



**Figure 1-2 : Extrait du schéma collocationnel pour la notion |DESCRIPTION| en français (Pecman, 2004, p. 357)**

Comme on peut le constater, sous la notion |DESCRIPTION| sont regroupées les différentes catégories verbales (*fournir une description*), nominales (*analyse détaillée*), adjectivales (*données descriptives*), etc. Le schéma permet d'aborder les unités phraséologiques renvoyant à une notion sous des formes diverses. Ces ressources nous semblent intéressantes du point de vue didactique puisqu'elles mettent à la disposition des apprenants différents moyens d'expression. D'après Pecman, les structures collocationnelles contribuent à la mise en place des différentes fonctions au sein d'un texte scientifique (voir la figure 1-3).

Ces ressources phraséologiques sont ensuite intégrées dans la base ARTES, accessibles par fonctions discursives. Dans la base, on compte 80 fonctions discursives telles que « annoncer les conclusions », « faire des illustrations ou donner des exemples », « souligner une différence, une dissemblance ou une opposition ». La figure 1-3 présente la fonction « faire un renvoi à ce qui vient d'être traité » :

The screenshot shows the 'DICTIONNAIRE ARTES' interface. The top navigation bar includes 'Terminologie en contexte', 'Phraséologie discursive', and 'Recherche multicritères'. The main content area is titled 'Fonctions discursives' and features a search result for 'comme évoqué précédemment'. On the left, there is a text box with a definition in French and an English translation. On the right, a 'Faire un renvoi à ce qui vient d'être traité' section contains a table of associated collocations. The table has three columns: 'Collocation', 'Construction', and 'Discours'. Below the table is a pagination control showing '1 / 1'.

Collocation	Construction	Discours
comme souligné dans la section X	construction verb	discours multirè
comme évoqué précédemment	construction verb	discours scienti
comme il a été mentionné	construction verb	discours multirè
comme nous l'avons vu dans X	introduceur d'év	discours scienti
comme expliqué dans	schéma phrastiqu	discours multirè
nonobstant ce qui précède	prép. N	discours univer
comme indiqué précédemment	introduceur d'év	discours scienti
comme indiqué dans	introduceur d'év	discours scienti
ceci suggère que	construction verb	discours multirè
précité	autre	discours iuridia

**Figure 1-3 : Accès aux unités phraséologiques par fonctions discursives dans la base ARTES**

Dans la partie droite, une dizaine de structures collocationnelles sont répertoriées sous cette fonction, par exemple, la construction en *comme* (*comme souligné dans la section X*, *comme évoqué précédemment*, *comme nous l'avons vu dans X*). Dans la partie gauche, certains exemples du corpus ainsi que la traduction (notamment vers l'anglais et le français) sont proposés. Néanmoins, il nous semble que le classement sémantique des structures collocationnelles pourrait être affiné à l'intérieur de chaque fonction. Par exemple, la structure *comme expliqué dans* renvoie à une explication alors que *comme nous l'avons vu dans X* relève d'un constat. Par ailleurs, certaines structures ne semblent pas remplir la fonction désignée, par exemple *ceci suggère que* implique plutôt une déduction ou un raisonnement. Nous pensons qu'il est utile de réfléchir sur un accès onomasiologique bien ciblé et organisé des données phraséologiques dans la conception d'un outil rédactionnel.

De manière générale, l'étude de Pecman et son application dans la base ARTES nous inspirent doublement. Du point de vue phraséologique, il nous semble important de croiser l'analyse des discours scientifiques et l'étude des phénomènes phraséologiques. Par ailleurs, à l'instar de Pecman (*ibid.*) et de Tutin (2010b), nous

pensons que l'étude de la phraséologie doit dépasser la forme syntaxique et prendre en compte les aspects sémantiques et discursifs. Ainsi, nous envisageons de proposer les classements des constructions verbales par notions sémantiques et par fonctions discursives dans notre analyse linguistique. Du point de vue lexicographique, la proposition d'un accès onomasiologique par fonction discursive, complémentaire à l'accès sémasiologique, nous semble pertinente dans la conception d'un outil d'aide à la rédaction scientifique.

#### **1.4.2. Intégrer les corpus dans l'élaboration et l'utilisation du dictionnaire : *The Louvain EAP Dictionary (LEAD)***

Suivant une approche structurelle fonctionnelle semblable à celle de l'*Academic keyword List (AKL)*, le *Louvain English for Academic Purposes Dictionary (LEAD)*, conçu par Granger et Paquot (2010 ; 2015), est un dictionnaire d'anglais à des fins académiques en ligne<sup>21</sup> visant à fournir aux apprenants un outil d'aide à l'écriture académique. Ce dictionnaire électronique est donc basé sur la liste AKL et décrit les éléments phraséologiques du lexique académique qui remplissent des fonctions importantes dans le discours académique.

La principale caractéristique de ce dictionnaire est qu'il est basé non seulement sur l'analyse de corpus d'experts (le corpus général *English for Academic Purposes corpus* et plusieurs corpus spécialisés en commerce, linguistique, médecine, littérature, éducation, etc.), mais aussi celle de corpus d'apprenants (le corpus *International Corpus of Learner English* et le corpus *Varieties of English for Specific Purposes Database*). Les corpus d'experts sont utilisés dans la description lexicographique des mots académiques. Sous chaque entrée, nous trouvons non seulement des informations lexicographiques comme la définition et les exemples issus du corpus mais aussi des syntagmes et des collocations autour de chaque entrée. La figure 1-4 illustre l'entrée du nom *source* dans le LEAD.

---

<sup>21</sup> L'accès du dictionnaire est réservé aux personnels et étudiants de l'Université catholique de Louvain : <http://130.104.6.153/test002/web/app.php/login>, consulté le 20 avril, 2017.

The Louvain EAP dictionary

Search Selected discipline: *Business*; selected mother tongue: *French* ([change settings](#))

Search by function      Search by translation      Corpus search      Exercises

**source** (n.)

the first or original cause of something or the place where it begins:

*In France, taxation remained a perennial source of tension between the crown and its subjects.*

*However, like questions about the sources of resources, future research should seek to **identify** the sources or causes of EO. (BUS)*

*This study extends this line of research by taking a real options approach, which enables us to **identify** various sources of uncertainty surrounding venture capital investments and to examine how exogenous and endogenous uncertainties influence the timing of staged financing. (BUS)*

*By developing a procedure based on the GT for **identifying** alternative sources of response variance in cross-national research, our research offers an alternative and/or a complementary method to CT for establishing cross-national validity of measurement scales. (BUS)*

*Moreover, this study applies the decomposition method to **identify** and compare sources of such compensation difference. (BUS)*

*For example, we **identify** the sources of stimuli on the basis of patterns of stimulation on the receptors. (EAP)*

**Collocations**

Adj + source  
*other, major, main, important, new, different, primary, secondary, only, various, possible, potential, alternative, single, significant, principal, useful*

V + source  
*provide, use, become, remain, identify, constitute, find, offer, consider, need, suggest, locate, represent, examine, tap, require, cite, trace, create, remove, ignore, determine, check*

source + V  
*provide, suggest, contribute, influence*

**Lexical bundles**

the source of  
 source of information  
 source material  
 source of data  
 a wide variety of sources

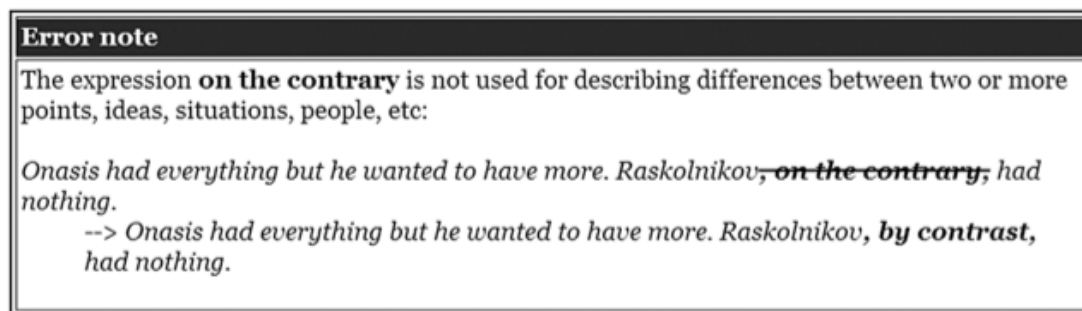
Figure 1-4 : Entrée du nom *source* dans le LEAD (Paquot, 2012, p. 175)

Comme on peut le constater, à droite de la figure sont proposés des syntagmes récurrents tels que *the source of*, *source of information* ainsi que des collocations du type « Adj+source » (*major, important, etc.*) et du type « V+source » (*find, identify, etc.*). Il est possible d'afficher des exemples de corpus pour chaque collocation ou syntagme. Il est intéressant de noter que les exemples affichés peuvent varier selon la discipline sélectionnée. Par exemple, dans la figure ci-dessus, on voit des exemples pour la collocation *identify source* relevant de la discipline du commerce.

De même que le dictionnaire ARTES, LEAD propose également un mode d'accès onomasiologique, basé sur une liste de 18 fonctions rhétoriques considérées comme primordiales dans les discours académiques dont la comparaison, l'illustration, l'exemplification, etc. À chaque fonction rhétorique est associée une liste d'expressions lexicales nominales, verbales, syntagmes, etc., qui aidera les apprenants à mieux s'exprimer. Par exemple, pour la fonction rhétorique de la comparaison, nous avons des expressions adverbiales comme *by comparison, by contrast, on the other hand, etc.*

Par ailleurs, l'élaboration du dictionnaire a pris en compte les données de corpus

d'apprenants afin d'identifier les difficultés et les besoins des apprenants. Ces derniers ont ainsi guidé la sélection des entrées, des éléments phraséologiques et des exercices ciblés<sup>22</sup>. Les erreurs relevées dans les corpus d'apprenants sont traitées avec des feedback sous forme de notes comme illustré dans la figure 1-5 :



**Figure 1-5 : Note d'erreurs pour l'entrée de l'expression *on the contrary* dans le LEAD (*ibid.*, p. 173)**

L'intégration des exercices et des feedbacks a fait de ce dictionnaire un outil centré sur l'apprenant. En fait, nous sommes du même avis que plusieurs chercheurs (Verlinde, Selva, & Binon, 2009 ; Granger & Paquot, 2015) à propos de l'élaboration d'un outil d'aide à la rédaction académique/scientifique destiné aux étudiants non-natifs qui doit être centrée sur les besoins réels des apprenants. Étant donné que le corpus d'apprenants donne un moyen d'identifier les besoins langagiers rencontrés par les apprenants, l'analyse de ces ressources nous paraît primordiale dans les études ayant des objectifs similaires. Ainsi, dans le cadre de notre étude, nous avons constitué un corpus d'apprenants (*cf.* Chapitre II) afin d'identifier les besoins liés à l'utilisation des verbes du LST et de leurs constructions verbales chez les apprenants (*cf.* Chapitre V).

Un autre point qui nous semble intéressant dans le dictionnaire LEAD est l'intégration dans le dictionnaire des corpus d'experts utilisés dans la description lexicographique. Ainsi, les apprenants peuvent accéder directement aux corpus d'experts depuis le dictionnaire en ligne, soit pour chercher un mot ou une expression, soit pour formuler des requêtes précises afin d'identifier des collocations ou des séquences récurrentes. D'après Paquot (2012), le fait de proposer une consultation

<sup>22</sup> Dans le dictionnaire, les exercices sont proposés en fonction des difficultés identifiées. Il s'agit notamment des exercices lacunaires, des exercices à choix multiples créés avec le logiciel Hot potatoes.

directe du corpus permet aux apprenants d'accéder à plus d'informations (ou à des informations plus ciblées) et de travailler en autonomie. On ne nie pas les avantages de la consultation directe du corpus par les apprenants. Par contre, cette proposition reste discutable, car les apprenants risquent d'être noyés par la masse d'informations dans le corpus. À l'instar de Chambers (2010) et Cavalla (Cavalla & Loiseau, 2013), nous pensons qu'au plan technique, les apprenants ont besoin d'être formés sur la manipulation du logiciel de concordance, par exemple, comment formuler une requête sur l'interface d'interrogation du corpus. Par ailleurs, comme le souligne Cavalla (*ibid.*, p. 166), du point de vue discursif, « la découverte des spécificités » liées aux éléments phraséologiques n'est pas évidente. Ainsi, nous rejoignons Cavalla et Chambers sur le fait qu'il est plus raisonnable de passer par une phase d'initiation où les apprenants peuvent travailler sur des lignes de concordances sélectionnées et des exercices préparés par les enseignants qui encouragent « une forme constructiviste d'apprentissage » (Chambers, 2010, p. 15). Le problème de la consultation directe des concordances est également soulevé par les auteurs eux-mêmes (Granger & Paquot, 2015). D'après ces derniers, l'outil doit aussi rendre les ressources du corpus plus accessibles aux non-linguistes. Cette idée est confirmée par les feedbacks des étudiants qui ne trouvent pas la consultation directe des corpus utile tandis qu'un groupe de doctorants et de chercheurs l'ont bien appréciée.

Nous reviendrons sur l'utilisation des corpus dans l'E/A des langues étrangères dans le chapitre II.

De manière générale, nous avons retenu plusieurs aspects intéressants dans le dictionnaire LEAD : l'intégration des corpus d'apprenants, les deux modes d'accès mettant en avant la phraséologie des mots de AKL, et la consultation des lignes de concordance.

### **1.4.3. Appliquer le CPA à l'écrit scientifique : *E-Advanced Learner's Dictionary of Verbes in Science (DicSci)***

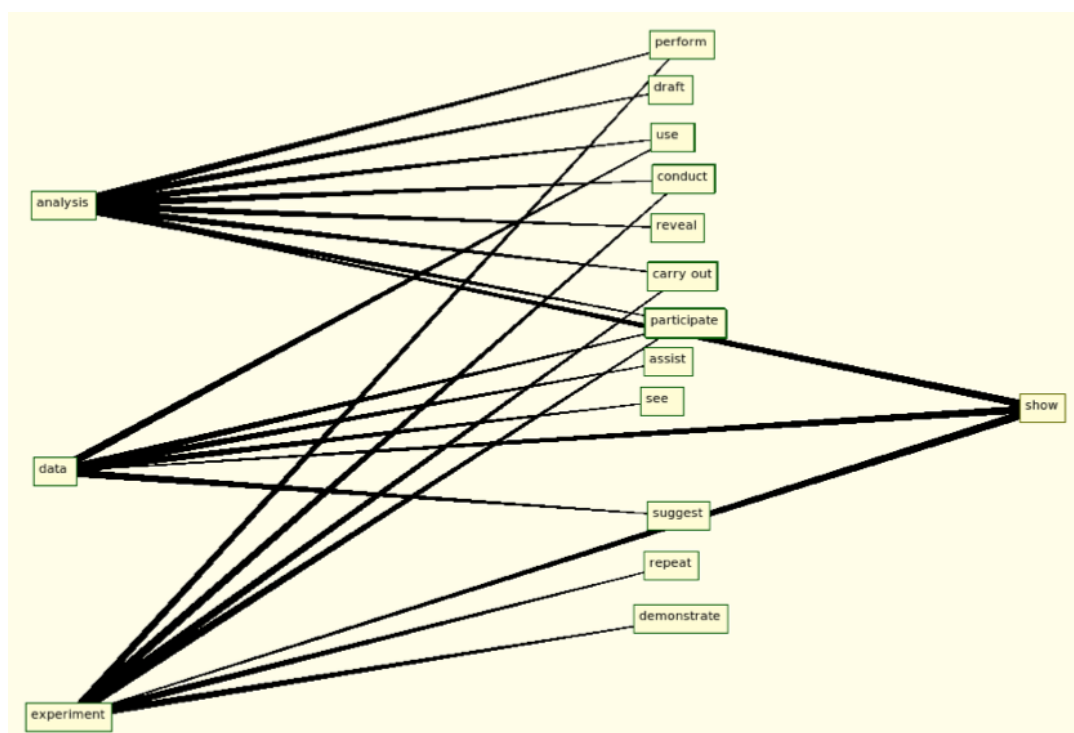
Comme nous l'avons mentionné dans la section 1.2.4., le travail de Williams et Million (2010 ; 2013) vise à appliquer la méthodologie CPA à l'analyse des verbes



dans l'écrit scientifique, ce qui débouche sur la construction d'un dictionnaire électronique spécialisé encore en cours d'élaboration. Ce dictionnaire nous intéresse particulièrement parce qu'il montre que la méthodologie CPA peut aussi être appliquée à cet objectif spécialisé qu'est l'écrit scientifique.

Comme les deux ressources mentionnées plus haut, le DicSi favorise aussi l'exploitation des éléments phraséologiques pour l'encodage de l'anglais. S'appuyant sur le corpus *BioMed Central corpus* (BMC), il est consacré à l'étude des collocations et des patrons lexico-syntaxiques autour du verbe et s'adresse à des locuteurs non-natifs, en particulier les biologistes et les médecins ayant besoin de rédiger en anglais.

La méthodologie adoptée se fonde sur trois éléments : les réseaux collocationnels, la résonance collocationnelle et les patrons lexico-syntaxiques du CPA (Williams, Millon, & Alonso, 2011). D'abord, à partir d'une liste des cent verbes les plus fréquents dans le corpus BMC, les auteurs ont étudié les réseaux collocationnels de ces verbes qui dirigent la sélection de nouvelles entrées lexicales. D'après Williams et Millon (2013, p. 49), le réseau collocationnel d'un mot donné est composé de ses co-occurents ainsi que des co-occurents de chaque co-occurent du mot donné. Par exemple, comme illustré dans la figure 1-6, le réseau collocationnel du verbe *show* est composé d'une série des co-occurents nominaux les plus significatifs du point de vue statistique dans le corpus comme *analysis*, *data* et *experiment*. En outre, à chaque co-occurent nominal (ex. *analysis*) s'ajoute une série des co-occurents verbaux les plus significatifs (*perform*, *use*, *conduct*, *reveal*, etc.).



**Figure 1-6 : Extrait du réseau collocationnel du verbe *show* (Williams & Millon, 2010, p. 1256)**

En fait, l'analyse des collocations vise d'une part à construire des classes sémantiques des verbes, c'est-à-dire que les verbes partageant un certain nombre de co-occurents sont susceptible d'être regroupés dans une même classe sémantique. D'autre part, cette analyse contribue à établir des patrons lexico-syntaxiques. Concernant le cas du verbe, ce sont notamment les co-occurents nominaux qui jouent un rôle important. Dans notre étude, l'identification des co-occurents du verbe constitue aussi une étape importante (*cf.* Chapitre III).

Ensuite, la résonance collocationnelle désigne « le processus qui montre les mouvements entre les acceptions en langue générale et celles en langue spécialisée » (*ibid.* p.50). Elle consiste donc à relier les sens « généraux » et « spécialisés » d'un mot. Dans le DicSci, la comparaison entre général et spécialisé s'appuie sur le *British National Corpus* (BNC) pour la langue générale et le BMC pour la langue spécialisée. Par exemple, comme le constatent les auteurs (*ibid.* p.60), un usage typique du verbe *probe* dans la langue générale – « utiliser une sonde afin d'examiner une blessure » – est transféré au sens « utiliser un instrument moléculaire afin d'étudier des aspects de l'organisme pour fournir des informations sur son état » dans le contexte de biologie

moléculaire, comme dans l'exemple suivant : *blot was probed with monoclonal antibodies* (La membrane est examinée avec des anticorps monoclonaux). Ainsi, du général vers le spécifique, la notion d'examen est constante, mais le type d'instrument utilisé change. L'étude de la résonance collocationnelle permet de mieux saisir le sens noyau du mot dans un contexte large. Par exemple, pour le verbe *probe*, le sens noyau identifié est « examiner ».

Ainsi, la notion de résonance collocationnelle (Williams, 2008b) présente un grand intérêt pour les études comme la nôtre qui cherchent à relier les usages de la langue générale à ceux de la langue spécialisée. Étant donné qu'il n'existe pas encore de ressource lexicographique spécialisée destinée à l'écrit scientifique du français, nous nous fondons sur la ressource lexicographique LVF – qui décrit les emplois des verbes dans la langue générale – pour construire les sens et les patrons identifiés dans notre corpus spécialisé de l'écrit scientifique. Dans notre étude, l'enjeu lié à l'étude de la résonance d'un mot donné est donc de repérer la notion sémantique constante d'un contexte à l'autre et en passant du général vers le spécifique, de spécifier le sens avec le comportement sémantique/syntaxique rencontré (ex. collocations, structures syntaxiques) dans l'écrit scientifique. Nous expliquerons plus en détail dans le chapitre III comment nous exploitons la résonance afin de relier les entrées verbales du LVF aux emplois spécifiques du corpus Scientext-SHS.

Enfin, l'analyse des patrons des verbes fondée sur le modèle CPA consiste à compléter les deux méthodes précédentes. Comme souligné plus haut, l'identification des patrons est basée sur l'observation du corpus BMC et l'extraction du réseau collocationnel. Nous illustrons dans la figure 1-7 les patrons identifiés pour le verbe *show* dans le dictionnaire.

- *EVIDENCE* patterns and collocations
  - [X] show an increase in [Y]
  - show (some/great/considerable) promise (71)
- *POSITION* patterns and collocations
  - As shown in figure X,
  - [X] is shown (schematically) in figure/table [N]
  - [X =tests] showed (no) significant differences in/between [comparison X & Y]
  - [X] shows similarity to|in [Y]

**Figure 1-7 : Patrons et collocations du verbe *show* dans le DicSi (Williams & Millon, 2010, p. 1255)**

Comme on peut le constater, les patrons identifiés ont permis de créer deux classes sémantiques auxquelles appartient le verbe *show*, soit les classes « evidence » et « position ». Les patrons lexico-syntaxiques mettent en avant des structures syntaxiques récurrentes telles que *as shown in figure X*.

De manière générale, ce dictionnaire nous semble intéressant dans la mesure où la méthodologie adoptée se rapproche de la nôtre. L'étude de la phraséologie est abordée sous différents angles : collocation, résonance collocationnelle et patrons lexico-syntaxiques. Enfin, il faut souligner que, le dictionnaire est destiné aux apprenants, mais que les auteurs n'en ont pas encore communiqué l'organisation des informations facilitant l'accès aux apprenants ainsi que la réflexion sur le plan didactique. De ce fait, il reste programmatique.

À l'issue de l'étude de trois ressources/outils à visée didactique portant sur les éléments phraséologiques, il nous semble important de préciser le type d'outil que nous souhaitons développer. Nous résumons ci-dessous les éléments que nous souhaitons mettre en avant dans cet outil :

- les patrons phraséologiques (inspirés du CPA) qui décrivent les emplois du verbe dans l'écrit scientifique
- les fonctions rhétoriques/discursives associées aux patrons
- les besoins langagiers ciblés
- l'accès à des exemples présélectionnés du corpus par les enseignants.

Dans le cadre de notre étude, nous nous limitons à aborder les données

linguistiques que nous souhaitons intégrer dans cet outil. Signalons que ces données linguistiques sont exploitables par les enseignants et les linguistes, mais ne s'adressent pas directement aux apprenants. Il faut donc envisager une étape de didactisation ultérieurement. Nous proposons des pistes didactiques (par exemple, comment exploiter les données) à la fin du chapitre V. Par ailleurs, notons que dans une étude précédente, nous avons déjà essayé d'intégrer quelques patrons verbaux dans l'outil Dicorpus<sup>23</sup> (Yan & Hatier, 2016) (*cf.* Chapitre V). En effet, l'outil que nous voulons développer pourrait prendre la forme de Dicorpus, ce dernier correspondant à notre objectif d'application pédagogique présenté ci-dessus, notamment dans la mesure où il permet à l'apprenant de consulter des concordances présélectionnées par l'enseignant. Il est encore en cours de construction (Tran & Falaise, 2016) et nous souhaitons l'enrichir ultérieurement avec notre échantillon de patrons.

Enfin, nous rejoignons l'idée de Tutin (2005) qu'un dictionnaire/un outil mettant les éléments phraséologiques spécialisés est indispensable pour l'E/A du français et peut constituer un complément aux ressources lexicographiques existantes comme le *Petit Robert*.

## 1.5. Bilan

Dans ce chapitre, nous avons d'abord présenté notre objet d'étude – les verbes du LST – et plus précisément les patrons des verbes du LST dans le domaine de l'écrit scientifique. En examinant plusieurs modèles consacrés à l'étude des constructions verbales (*Grammaires de Construction*) ou aux patrons verbaux (*Pattern Grammar*), nous avons adopté le modèle CPA de Hanks pour étudier les propriétés sémantiques et syntaxiques des verbes du LST. Ce modèle fondé sur des usages réguliers des verbes, permet d'accéder aux sens du verbe à travers l'analyse des constructions syntaxiques et des co-occurrences associées du verbe.

---

<sup>23</sup> Outil, conçu par Achille Falaise (docteur en informatique et ingénieur de recherche) et Thi Thu Hoai Tran (docteure en FLE), fait office d'assistant phraséologique pour la rédaction scientifique en français et constitue une adaptation didactique du corpus Scientext : <http://corpora.aiakide.net/scientext18/services/dicorpus/doc/>

Si nous nous intéressons à l'analyse des patrons, c'est parce que nous sommes du même avis que Hanks et d'autres linguistes (comme les distributionnalistes ou les contextualistes), pour lesquels le mot n'a pas de sens intrinsèque, mais des sens potentiels qui peuvent être activés par les mots qui l'entourent. En ce sens, les patrons lexico-syntaxiques liés aux mots sont révélateurs des sens du mot, et cette démarche est rendue possible par l'étude de corpus.

À travers la présentation des travaux autour de la méthodologie CPA, nous avons montré que ce modèle peut être appliqué à différentes études ayant un objectif linguistique ou didactique, dans le domaine de la langue générale ou de la langue spécialisée. En ce qui concerne notre étude, nous pensons que le modèle CPA peut être appliqué à l'étude sur l'écrit scientifique du français, en permettant de bien mettre en évidence les propriétés sémantiques et syntaxiques des verbes du LST.

Sur le plan didactique, nous avons souligné que la phraséologie du LST constitue des difficultés majeures pour les apprenants. Afin de répondre à ces problèmes, nous avons mené nos premières réflexions en vue des applications pédagogiques sur les patrons verbaux, notamment en examinant trois ressources phraséologiques qui inspirent notre conception d'un outil d'aide à la rédaction scientifique (*cf.* Chapitre V).

Nous allons passer maintenant au chapitre II où nous présentons les données utilisées dans notre étude.

---

## Chapitre II.

# Constitution et exploration des corpus

---

2.1. UNE APPROCHE INSCRITE DANS LA LINGUISTIQUE DE CORPUS.....	61
2.1.1. La linguistique de corpus : apports et limites.....	62
2.1.2. Une approche associant les analyses « corpus-based » et « corpus-driven » .....	63
2.2. L'UTILISATION DE TROIS CORPUS DANS NOTRE ANALYSE LINGUISTIQUE AVEC UN OBJECTIF DIDACTIQUE PRECIS .....	64
2.2.1. Les usages de trois corpus dans notre étude .....	64
2.2.2. L'intérêt des corpus dans l'enseignement et l'apprentissage des langues étrangères .....	66
2.3. LES CORPUS D'ANALYSE.....	68
2.3.1. Le corpus d'analyse : Scientext-SHS.....	68
2.3.2. Le corpus d'apprenants : corpus Sup-Chinois .....	70
2.3.3. Le corpus d'étudiants natifs : corpus Sup-Natifs .....	76
2.3.4. Les choix des méthodes d'analyse pour le diagnostic des besoins langagiers.....	78
2.3.5. Le tableau récapitulatif de comparabilité.....	80
2.4. LES OUTILS D'EXPLOITATION .....	81
2.4.1. ScienQuest .....	82
2.4.2. Le Lexicoscope .....	86
2.4.3. TXM .....	89
2.5. BILAN .....	91

---

Dans ce chapitre, nous présentons les données de corpus utilisées dans notre analyse linguistique avec un objectif didactique précis – l'enseignement des constructions verbales du LST pour aider à la rédaction académique. L'analyse linguistique se compose de deux parties, la première relative à l'étude des patrons lexico-syntaxiques des verbes du LST dans l'écrit scientifique (*cf.* Chapitre IV), la seconde à l'analyse des besoins langagiers liée à l'utilisation des verbes du LST chez les apprenants chinois (*cf.* Chapitre V). Nous nous posons principalement les questions suivantes : quel corpus de référence permet l'étude des caractéristiques linguistiques de l'écrit scientifique ? Quelles données collecter pour identifier les difficultés spécifiques concernant l'emploi des verbes et des constructions verbales du LST pour les apprenants chinois et les scripteurs français novices dans l'écriture académique ?

Pour ce faire, nous avons utilisé trois corpus. D'abord, le corpus d'experts Scientext-SHS, utilisé en amont pour relever des constructions verbales caractéristiques des pratiques expertes et en aval comme une source de données pour

l'exploitation didactique. Ensuite, le corpus Sup-Chinois, que nous avons constitué pour relever les erreurs ainsi que les cas de sous-emploi et de suremploi liés à l'utilisation des verbes et des constructions verbales du LST chez les apprenants chinois. Enfin, le corpus Sup-Natifs a servi à mener une comparaison mettant en lumière les difficultés des apprenants chinois.

Nous commencerons par justifier le choix de l'approche sur corpus adoptée pour l'étude des patrons verbaux. Ensuite, après avoir présenté la constitution des trois corpus, nous nous interrogerons sur les questions de genre et de comparabilité liées aux corpus. Pour finir, nous abordons les outils d'exploitation qui nous sont utiles pour l'analyse sur corpus.



## 2.1. Une approche inscrite dans la linguistique de corpus

La linguistique de corpus issue de la tradition anglo-saxonne, a été initialement développée pour concevoir des dictionnaires d'apprentissage<sup>1</sup> destinés aux apprenants de l'anglais langue étrangère. De ce fait, comme le fait remarquer Williams (2005), dès ses débuts, la linguistique de corpus est étroitement liée à la lexicographie pédagogique et à l'enseignement des langues (étrangères). C'est en fait au croisement de ces trois disciplines que s'inscrit notre étude.

Considérée comme «à fois une discipline à part et une méthodologie sur corpus» (*ibid.*, p. 13), la linguistique de corpus permet d'étudier «la langue en contexte sous la forme de grands ensembles de textes, les corpus» (*ibid.*, p. 13). Ici, le corpus est défini comme un groupement de données structurées en respectant les objectifs théoriques et pratiques de recherche, allant du choix des discours et des genres aux applications (Rastier, 2002).

Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications. (Rastier, 2002, p. 2)

Aujourd'hui, on voit l'apparition de corpus et d'approches sur corpus diverses, en fonction de différents objectifs et pratiques de recherche. Dans notre étude, nous prenons pour corpus d'analyse différents ensembles de textes se rapportant aux genres de l'écrit scientifique et académique : corpus d'articles scientifiques, corpus d'écrits universitaires d'étudiants chinois (corpus Sup-Chinois) et corpus d'écrits universitaires d'étudiants français (corpus Sup-Natifs). Avant de présenter les différents objectifs de recherche liés à la constitution de ces corpus, il nous semble important d'expliquer notre position par rapport à l'utilisation de la linguistique de corpus pour la description de la langue et pour son enseignement à des publics allophones.

---

<sup>1</sup> Nous pouvons citer par exemple le dictionnaire *Idiomatic and Syntactic English* édité (ISED) par A.S. Hornby et le dictionnaire *Collins Cobuild English Dictionary* (COBUILD) édité par J. M. Sinclair.

### 2.1.1. La linguistique de corpus : apports et limites

Aujourd'hui, la linguistique de corpus s'avère incontournable pour étudier la langue en contexte, notamment les phénomènes phraséologiques. À cet égard, nous pouvons retenir deux aspects. Au niveau des données, d'une part, il est possible d'accéder à une grande quantité de données réelles pour observer et décrire l'emploi des mots tel qu'il est pratiqué dans le genre du corpus. Sur le plan phraséologique, le corpus donne les moyens de mettre en évidence les fréquences et les régularités de la langue telles que les associations des mots, les collocations et les patrons en termes de fréquence. Au niveau des outils d'analyse, d'autre part, la linguistique de corpus reliée aux techniques du traitement automatique des langues (TAL) fournit des outils informatiques permettant de traiter de nombreuses données de manière rapide et plus objective. Ainsi, nous déduisons deux caractéristiques principales de la linguistique de corpus, à savoir « l'objectivité et la représentativité » (Cori & David, 2008, p. 115), ce qui la distingue des approches basées sur l'intuition des linguistes.

Néanmoins, il importe de souligner que les valeurs représentative et objective liées à la linguistique de corpus ne sont pas absolues du fait que les données du corpus résultent d'une sélection (Jacques, 2005, p. 26). Ici, nous rejoignons la réflexion présentée par plusieurs chercheurs (Sánchez Cárdenas, 2010 ; Cori & David, 2008) selon laquelle le corpus constitue simplement une aide au processus de recherche et la linguistique de corpus n'efface pas les difficultés rencontrées dans la recherche scientifique, puisque l'interprétation et la vérification des résultats sont indispensables. Nous pensons qu'un corpus, quelle que soit sa taille, ne peut jamais représenter tous les usages du mot. C'est le linguiste qui joue un rôle important pour « décider ce qui est significatif dans ce qu'il observe » (*ibid.*, p. 120). Rappelons que les faits scientifiques s'analysent à partir de l'interprétation des données et non pas par la simple observation. Ainsi, dans notre étude, nous n'envisageons pas de relever tous les emplois des verbes du LST présents dans le corpus Scientext-SHS, mais nous nous focalisons sur les constructions verbales du LST les plus fréquentes et sur l'analyse de leurs propriétés sémantiques et syntaxiques dans le contexte scientifique.

Par ailleurs, comme le fait remarquer Sánchez Cárdenas (2010), si l'on considère

les outils informatiques du TAL associés à la recherche sur corpus, le traitement automatique des données est seulement un support et il ne peut pas remplacer la réflexion du linguiste, notamment en matière d'interprétation sémantique. Par exemple, nous avons opté pour un repérage semi-automatique des patrons verbaux dans le chapitre III. Il s'agit d'un repérage associant l'extraction automatique des constructions verbales ainsi que la vérification et l'analyse manuelles des patrons verbaux. Le chercheur doit considérer ces deux étapes indispensables, car les techniques de TAL ne peuvent pas désambiguïser automatiquement les sens du verbe.

Arrêtons-nous maintenant sur l'approche sur corpus adoptée pour l'étude des patrons verbaux.

### **2.1.2. Une approche associant les analyses « corpus-based » et « corpus-driven »**

Dans le domaine de la linguistique de corpus, on peut distinguer deux types d'approches sur corpus, à savoir l'approche *corpus-based* (basée sur le corpus) et l'approche *corpus-driven* (guidée par le corpus) (Tognini-Bonelli, 2001). Dans la première approche – *corpus-based* –, les corpus sont utilisés pour « confirmer au moyen de données authentiques l'analyse menée par le linguiste » (Azzopardi, 2010, p. 13). Cette approche s'appuie sur des théories linguistiques préalables et le corpus aide à les confirmer ou non. La deuxième approche – *corpus-driven* – constitue « une démarche inductive partant du corpus pour aller vers une éventuelle modélisation des régularités émergentes des données authentiques traitées » (*ibid.*, p. 13). Dans ce cas, l'unique hypothèse de départ est que le corpus doit révéler des éléments à analyser ensuite. Donc le linguiste n'exploite pas le corpus pour vérifier un fait, mais pour faire émerger des phénomènes non envisagés jusque-là.

Pour notre part, nous avons choisi de combiner ces deux approches qui nous semblent compatibles et complémentaires. D'une part, nous privilégions l'approche *corpus-based* dans la mesure où les corpus sont d'abord utilisés pour valider nos hypothèses et illustrer le modèle linguistique développé. Par exemple, l'analyse des deux corpus d'écrits d'étudiants permet de vérifier l'hypothèse selon laquelle

manipuler des constructions verbales dans des écrits académiques est plus difficile pour des apprenants chinois que pour des étudiants natifs du français. En outre, étant donné que nous adoptons la ressource lexicographique du LVF comme point de départ de notre analyse linguistique portant sur les verbes du LST dans le corpus Scientext-SHS, nous partons du point de vue selon lequel il y a adéquation entre les constructions syntaxiques et leur interprétation sémantique (Dubois & Dubois-Charlier, 1997) (*cf.* Chapitre I). Ce postulat théorique est important pour guider notre analyse linguistique, permettant de proposer une classification sémantique des verbes du LST (*cf.* Chapitre III)

D'autre part, nous nous situons dans la lignée d'une linguistique « corpus-driven » en ce que nous nous inspirons du modèle CPA pour analyser les propriétés sémantiques et syntaxiques des verbes du LST et proposer une modélisation des patrons verbaux. Rappelons que le CPA est une approche guidée par le corpus (*cf.* Chapitre I). Nous nous intéressons à la description des emplois du verbe et les phénomènes phraséologiques tels qu'ils sont pratiqués dans l'écrit scientifique parce que l'objectif principal de notre recherche est de proposer une ressource linguistique que sont les constructions verbales trouvées dans les écrits scientifiques des experts, donc utiles aux non-natifs dans la rédaction de leurs écrits académiques.

Dans la section suivante, nous aborderons les usages des trois corpus mentionnés ci-dessus dans notre analyse linguistique avec un objectif didactique précis – l'enseignement des constructions verbales du LST pour aider à la rédaction académique.

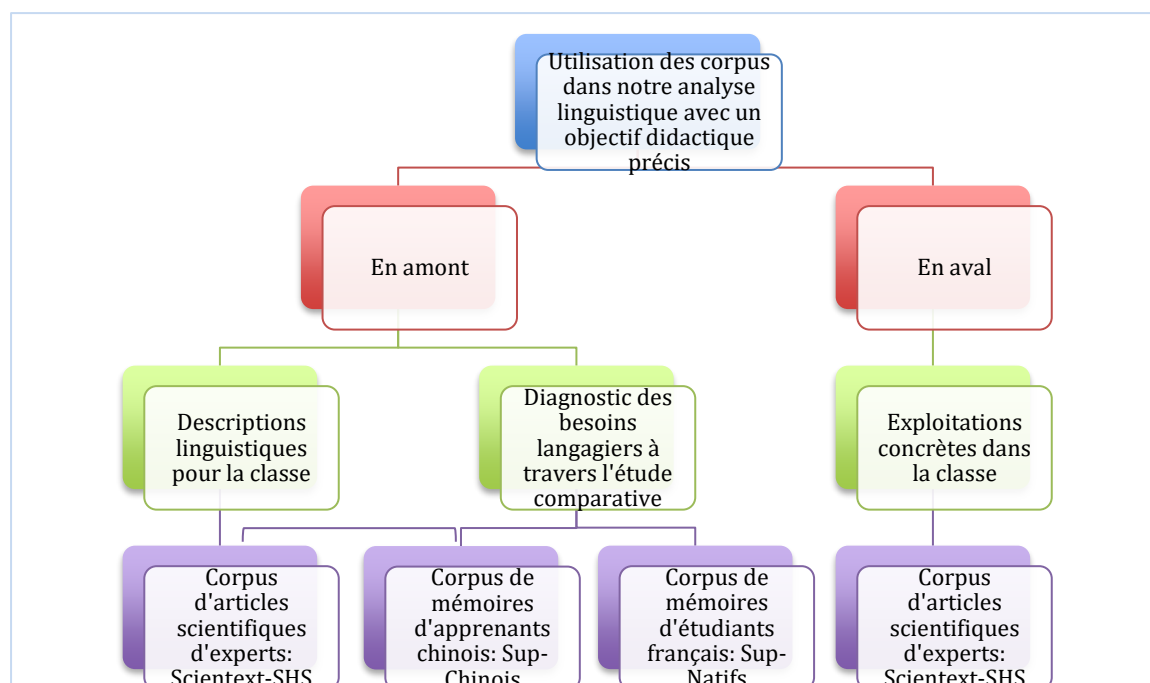
## **2.2. L'utilisation de trois corpus dans notre analyse linguistique avec un objectif didactique précis**

### **2.2.1. Les usages de trois corpus dans notre étude**

Dans le cadre du projet Scientext, plusieurs chercheurs ont utilisé le corpus d'écrits scientifiques Scientext afin d'élaborer des séquences didactiques pour aider la rédaction académique en français (Montagnon 2007; Tran, 2009, 2014 ; Cavalla,

2015). Comme ces travaux, notre intérêt porte sur l'étude des phénomènes phraséologiques dans une perspective didactique. Cependant, notre réflexion se distingue de ces dernières sur le point suivant : nous cherchons à délimiter les objets d'enseignement à partir des difficultés observés dans les corpus d'apprenants. À notre connaissance, il existe de nombreuses études consacrées à la description des phénomènes phraséologiques de la langue, mais peu d'entre elles se centrent sur les difficultés des apprenants en français comme langue étrangère. Notons que nous suivons ces études à propos de l'enseignement de ces éléments à partir des données du corpus (*cf.* Chapitre V).

Enfin, notre étude vise à montrer que les corpus spécialisés comme Scientext peuvent contribuer à l'amélioration de l'E/A dans le domaine de l'écrit scientifique, et cela tant du point de vue linguistique, pour cibler les contenus d'enseignement, que du point de vue didactique, pour intégrer les données authentiques, dans la classe de FLE notamment. En nous inspirant de la proposition de Römer (2010)<sup>2</sup>, nous illustrons les usages des trois corpus constitués dans le cadre de notre étude dans la figure 2-1 :



**Figure 2-1 : Les usages des trois corpus constitués dans notre étude**

<sup>2</sup> Cette figure se base sur la figure 2.1 proposée par Römer (2010, p. 19), elle constitue une version adaptée à notre cas d'étude.

Comme indiqué dans la figure 2-1, dans le domaine spécialisé de l'écrit scientifique, les trois corpus sur lesquels s'appuie notre étude seront exploités différemment. En amont de l'étude, nous exploitons d'une part le corpus d'experts Scientext-SHS afin de décrire le comportement lexico-syntaxique des verbes du LST. Nous souhaitons ainsi relever les emplois des verbes du LST les plus fréquents que les experts ont tendance à produire de manière préférentielle et répétitive dans l'écrit scientifique. D'autre part, en comparant le corpus Sup-Chinois aux corpus Scientext-SHS et Sup-Natifs, nous visons à établir un diagnostic lié à l'utilisation des verbes du LST chez les apprenants chinois par rapport aux scripteurs natifs. Nous envisageons de mettre en évidence les difficultés dans l'usage de la phraséologie scientifique chez des apprenants à travers l'analyse des erreurs et des cas de sous-emplois (*cf.* Chapitre V). Les éléments linguistiques identifiés dans l'analyse devraient faire l'objet d'un enseignement systématique de l'écrit.

En aval de l'étude, les exemples authentiques du corpus Scientext-SHS ainsi que les données issues de notre analyse seront intégrés dans les pratiques enseignantes. En effet, les résultats de notre recherche linguistique présentés sous forme de patrons verbaux mettent en évidence les associations des mots ainsi que les constructions syntaxiques récurrentes dans l'écrit scientifique. Ces ressources, mises à la disposition des enseignants, permettent de proposer des exercices et éventuellement donnent des pistes didactiques pour l'E/A de l'écrit académique.

Après avoir précisé les différents usages de trois corpus dans le cadre de notre étude, nous présenterons dans la section suivante l'intérêt des corpus dans l'E/A des langues étrangères.

### **2.2.2. L'intérêt des corpus dans l'enseignement et l'apprentissage des langues étrangères**

La linguistique de corpus est étroitement associée à l'E/A des langues étrangères, comme le montrent les projets pionniers de dictionnaires pour apprenants COBUILD de Sinclair (1987, 2004) ou les travaux de la recherche en « data-driven learning » (DDL) initiés par Johns (1991). De manière générale, l'exploitation des corpus dans

ce domaine peut être envisagée sous deux aspects différents.

D'une part, au niveau de la recherche linguistique à visée didactique (Römer & Wulff, 2010 ; Tutin & Grossmann, 2014), le corpus permet de décrire les usages pratiqués par les scripteurs natifs « pour pouvoir servir de modèle aux outils pédagogiques. » (Granger, Vandeventer, & Hamel, 2001, p. 610). Il est communément admis que l'objectif premier du corpus est de fournir des extractions permettant une description du fonctionnement de la langue en contexte. En fait, dans notre étude, le corpus d'experts Scientext-SHS contribue à révéler ce que Hanks (2000) appelle la « linguistic behaviour » des verbes, c'est-à-dire, « l'environnement syntagmatique » des verbes (les collocations ou les combinaisons fréquentes des mots) et « les relations entre cet environnement syntagmatique et le sens » des verbes qui les composent (Béjoint, 2007, p. 17). Nous rejoignons Paquot (2010) qui propose de prendre en compte cinq types principaux d'informations dans une étude ayant pour objectif didactique l'aide à l'écrit scientifique : propriétés sémantiques, distributions syntaxiques, collocations, fréquences et différences de registre. Nous pensons que seule une analyse du corpus permet de fournir ces informations.

Par ailleurs, il faut aussi mentionner les corpus d'apprenants, souvent constitués dans le but de prendre en compte les productions d'apprenants en cours d'acquisition. Ils peuvent donner accès aux erreurs et aux difficultés des apprenants (Boulton et al., 2013). En outre, la comparaison des corpus d'apprenants par rapport aux corpus de natifs permet de révéler les utilisations spécifiques liées aux apprenants (ex. : le sous-emploi ou le suremploi des mots, des phrases et des constructions) (Paquot, 2010) et d'aider à cibler des éléments linguistiques précis qui sont absents des discours des apprenants. Ainsi, les corpus d'apprenants répondent aux objets/objectifs d'enseignement puisque les résultats du diagnostic à partir des corpus d'apprenants peuvent être intégrés dans les dictionnaires d'apprentissage, les manuels, les méthodes ou les outils d'aide à la rédaction, etc. (Gilquin et al., 2007)

Le corpus d'apprenants Sup-Chinois utilisé dans notre étude a été constitué auprès d'étudiants chinois en master de français à l'Université des Études Internationales de Xi'an. Nous l'avons comparé avec un corpus d'écrits d'étudiants français en master didactique du français à l'Université Grenoble Alpes. Nous

pensons que ces deux corpus de petite taille ne peuvent représenter que des situations d'enseignement précises. En revanche, même s'ils sont dotés d'une valeur représentative moins importante, ils nous fournissent «une plus grande facilité d'analyse, de prise en main, d'interprétation et de constitution » (Aston, 1997 ; cité par Cavalla & Loiseau, 2014, p. 168), permettant d'établir un diagnostic plus ciblé et de créer des matériaux didactiques plus spécifiques.

D'autre part, au niveau de l'enseignement et de l'apprentissage, les corpus se révèlent efficaces par apport à d'autres formes de pratiques, et ceci dans différents contextes d'enseignement (Cobb & Boulton, 2015). Nous reviendrons sur cet aspect dans la section 5.5 du chapitre V.

Nous abordons dans la section suivante les différentes étapes de constitution des trois corpus.

## **2.3. Les corpus d'analyse**

### **2.3.1. Le corpus d'analyse : Scientext-SHS**

Notre corpus d'analyse, Scientext-SHS<sup>3</sup>, est utilisé non seulement pour la description linguistique, mais aussi pour l'exploitation didactique. Ce corpus, d'environ 5 millions de mots français, est composé de 500 articles scientifiques des 10 disciplines de sciences humaines et sociales suivantes : linguistique, psychologie, sciences de l'éducation, économie, sciences politiques, anthropologie, histoire, géographie, sciences de l'information et de la communication [SCI], sociologie. Ayant pour objectif d'« étudier le positionnement et le raisonnement de l'auteur dans les écrits scientifiques » (Tutin *et al.*, 2009), les linguistes ont constitué ce corpus dans le cadre du projet ANR Scientext. Il a été ensuite complété par l'ANR TermITH par Tran (2014), puis Hatier et Jacques (Hatier, 2013 ; Hatier *et al.*, 2014). Les sous-corpus correspondant à chaque discipline recensent chacun 50 articles. La

---

<sup>3</sup> Parmi les 500 articles scientifiques qui composent notre corpus d'analyse, seuls 164 articles, d'environ 1 million de mots, relevant de quatre disciplines (linguistique, psychologie, économie, sciences du langage) sont diffusables et consultables en ligne.



diversité disciplinaire et l'homogénéité du genre de l'article nous permettent d'extraire le lexique transversal aux disciplines. Le tableau 2.1 indique le nombre de mots de chaque sous-corpus :

Discipline	Nombre d'articles	Nombre de mots
Anthropologie	50	493 988
Économie	50	417 944
Géographie	50	400 533
Histoire	50	773 170
Linguistique	50	425 952
Psychologie	50	417 846
Sciences de l'éducation	50	417 069
Sciences de l'information	50	399 007
Sciences politiques	50	548 222
Sociologie	50	540 630
<b>Total Corpus Analyse</b>	<b>500</b>	<b>4 834 361</b>

**Tableau 2.1 : Composition du corpus d'analyse**

Le genre choisi est celui de l'article de recherche, ce qui s'explique notamment par son statut emblématique – « indicateur majeur de la production des chercheurs et de la visibilité de leurs travaux » – permettant de caractériser les communautés de discours dans l'activité scientifique (Rinck, 2010a, p. 436). Au sein des variations disciplinaires, l'article de recherche « obéit à des règles et à des codes particuliers au niveau du contenu et de la forme » (Boure, 1998 : 107, cité par Poudat, 2006, p. 51). Il se caractérise, d'une part, par les dimensions énonciative et rhétorique visant à répondre à une problématique par des argumentations en faveur du point de vue de l'auteur tout en essayant de rester objectif ; d'autre part, par une structure textuelle homogène divisée en sections et séquences identifiables : « l'introduction (contexte de recherche, etc.), le développement (organisé ou non selon une structure IMRaD), la conclusion (perspectives, bilan, etc.) » (Tran, 2014, p. 56).

### 2.3.1.1. Les annotations du corpus

Ces articles proviennent de numéros récents de revues avec comité de relecture, dont l'évaluation par l'AERES<sup>4</sup> est A ou B. Ce corpus de 5 millions de mots a été balisé en parties textuelles (introduction, conclusion, notes de bas de page, etc.), et mis au format XML-TEI (Falaise, 2014). Par ailleurs, il a été annoté aux niveaux morphologique et syntaxique à l'aide de l'analyseur en dépendances XIP, un analyseur qui vise à extraire des dépendances syntaxiques de façon robuste (Aït-Mokhtar, Chanod, & Roux, 2002) : dans un premier temps, chaque lexème (mots, ou expressions figées, ex. : *analysé*) est étiqueté selon la forme (ex. : *analysé*), le lemme (ex. : *analyser*), la partie du discours (catégorie syntaxique, ex. : verbe) et les traits flexionnels (genre, nombre, temps, par exemple le participe passé).

Dans un deuxième temps, les annotations syntaxiques en dépendances consistent à relier les lexèmes par les relations syntaxiques, ce qui permet d'extraire les éléments linguistiques comme les collocations et les patrons (*cf.* Chapitre III). Soulignons que cette annotation syntaxique est particulièrement pertinente pour notre étude des constructions verbales.

Le corpus est ainsi enrichi en lemmes, catégories syntaxiques, traits morphosyntaxiques et relations syntaxiques. Il faut noter que le travail des annotations linguistiques a été réalisé automatiquement (Hatier et *al.*, 2016).

### 2.3.2. Le corpus d'apprenants : corpus Sup-Chinois

#### 2.3.2.1. Les objectifs relatifs à la constitution d'un corpus d'apprenants

Il est communément admis que la langue de l'apprenant appelée aussi « l'interlangue » (Galligani, 2003) se caractérise non seulement par les erreurs, mais aussi par « des fréquences de mots, d'expressions et de structures très différentes, certains éléments étant sur-employés et d'autres considérablement sous-employés » (Granger, 2007, p. 117). Les corpus d'apprenants permettent d'« observer des

---

<sup>4</sup> Agence d'évaluation de la recherche et de l'enseignement supérieur. Depuis 2014, cette agence est devenue l'HCERES : Haut Conseil de l'Évaluation de la Recherche et de l'Enseignement Supérieur.

régularités d'utilisation et de quantifier certains aspects du développement langagier et de l'interlangue en général » (De Cock et Tyne, 2014, p. 139).

Ainsi, afin d'identifier les difficultés des apprenants liées à l'utilisation des verbes et des constructions verbales du LST pour les aider ensuite dans leurs écrits académiques, il nous semble important de prendre en compte à la fois les erreurs ainsi que les cas de sous-emploi et de suremploi. Pour répondre à ce besoin, nous avons donc constitué un corpus de FLE des non-natifs (le corpus Sup-Chinois) qui permet de relever les erreurs centrées sur les verbes du LST dans les écrits académiques des apprenants. Par ailleurs, nous avons utilisé un corpus de FLM des étudiants natifs (le corpus Sup-Natifs) qui permet de mettre en évidence les cas de sous-emploi et de suremploi liés à l'utilisation des verbes et des constructions verbales du LST chez les apprenants chinois.

Dans l'étude comparative, nous postulons que le statut de novice en matière de rédaction scientifique confronte identiquement les étudiants natifs et non-natifs aux difficultés d'usage de la phraséologie scientifique. Nous nous situons donc dans la perspective de l'apprentissage des verbes et des constructions verbales du LST par des scripteurs novices et envisageons d'étudier leur usage par une comparaison des productions d'étudiants natifs et non-natifs et de textes de chercheurs, considérés ici comme experts de l'écriture scientifique.

### **2.3.2.2. Les critères de choix concernant le corpus d'apprenants Sup-Chinois**

Nous avons collecté notre corpus d'apprenants à l'Université des Études Internationales de Xi'an. Le choix de l'établissement peut s'expliquer par plusieurs raisons : tout d'abord, j'avais suivi un cursus de licence en spécialité de français à cette université. J'ai donc un contact personnel avec le département de Langue et Littérature Françaises où j'ai proposé un enseignement de français d'un semestre (de septembre à décembre 2014) et ma première expérimentation didactique sur l'enseignement des constructions verbales du LST. Ensuite, dans le département de Langue et Littérature Françaises, lors des pratiques rédactionnelles de leur mémoire de fin d'études, les étudiants de master se trouvent confrontés à des difficultés langagières majeures en raison du manque d'outils appropriés et de formation

spécifique à l'écrit académique. En effet, on constate d'une part un manque d'entraînement à la rédaction de textes académiques, et d'autre part l'utilisation des dictionnaires inadaptés (le *Larousse bilingue*, le dictionnaire *Frhelper*<sup>5</sup> et des dictionnaires électroniques sino-français qui proposent des traductions des mots isolés en chinois sans prendre en compte le contexte) pour une tâche de production écrite en français.

Les données recueillies sont composées de 28 mémoires de master<sup>6</sup> en spécialité de français (environ 600 000 mots). Leur sujet d'étude varie et porte soit sur la traduction (9), la littérature (12), la linguistique (1) ou des thématiques socioculturelles (6). Lors de la rédaction du mémoire, les étudiants (entre 24-26 ans) ont suivi en moyenne six ans d'études de français. Les évaluations et les notes sur les mémoires révèlent que le niveau en langue des étudiants chinois est B2 ou C1 (et la plupart B2) selon les niveaux du CECRL. Il s'agit donc d'un corpus d'apprenants homogène. Soulignons que ces étudiants n'ont jamais appris le français en immersion. L'évaluation des mémoires porte sur quatre aspects :

- le sujet du mémoire
- l'argumentation autour du sujet choisi
- la structure du mémoire
- la qualité de la langue

Ainsi, la rédaction des mémoires doit répondre non seulement aux exigences de la langue (orthographe, lexique et syntaxe), mais aussi aux exigences argumentatives et discursives. Autrement dit, l'étudiant doit montrer à travers la rédaction non seulement ses connaissances sur la langue, mais aussi exprimer son point de vue, formuler des hypothèses, synthétiser, établir la cause et l'effet, illustrer avec des exemples et construire une argumentation, etc.

Il faut noter que les mémoires des apprenants chinois ne sont pas complètement

---

<sup>5</sup> Un dictionnaire bilingue (chinois-français) en ligne : <http://www.frdic.com/>

<sup>6</sup> Le mémoire de Master (硕士论文 *shuòshì lùnwén* en chinois) des promotions 2010 et 2011 comporte environ 60 pages hors annexes. Il est rédigé pendant les deux dernières années du cycle de 3 ans.

authentiques, car ils sont partiellement corrigés par les enseignants chinois, mais ceci n'est pas gênant, sachant que les corrections ne portent que sur les erreurs grammaticales (en particulier l'orthographe) et plus rarement sur les erreurs lexicales ou syntaxiques. Par exemple, on peut trouver des erreurs lexicales comme *s'acharner à décrire la différence* ou des erreurs sur les constructions comme *le processus peut illustré comme*. Par ailleurs, de nombreux passages incompréhensibles nous amènent à penser que la plupart des mémoires ne sont pas complètement corrigés, comme en témoigne l'exemple suivant :

Considérer que le passé simple présente le procès dans sa globalité, en expliquant notamment son incompatibilité avec depuis + indication ponctuelle ou depuis + durée quantifiée, qui « suppose au contraire que le phénomène soit évalué dans sa durée à partir d'une saisie à un moment donné de son accomplissement ». (apprenants chinois\_linguistique\_1)

Dans cet exemple, il est difficile de repérer le sujet de la phrase. Les structures syntaxiques erronées rendent le texte incompréhensible. D'ailleurs, nous pouvons avancer que les enseignants non-natifs ne connaissent pas toujours très bien les éléments phraséologiques de la langue et que donc on pourra y retrouver des erreurs inconnues de ces enseignants. En ce sens, les données collectées constituent des indicateurs lexicaux qui fixent notre attention sur les erreurs effectives liées aux constructions verbales et sur les besoins lexicaux ainsi manifestés.

### 2.3.2.3. La définition de corpus d'apprenants

Par rapport aux corpus des locuteurs natifs, les corpus d'apprenants représentent les variétés des non-natifs dans l'apprentissage d'une langue seconde ou étrangère. En ce qui concerne le nôtre, il s'agit d'un corpus qui représente une variété des étudiants chinois dans l'apprentissage de français langue étrangère. À l'instar de Gilquin, nous retenons deux aspects essentiels dans la définition des corpus d'apprenants :

- Naturel : les corpus d'apprenants peuvent présenter différents degrés de naturalité selon différents contextes de production (Granger, 2008). L'idée est d'exclure les productions effectuées dans des situations imaginaires avec des messages non authentiques pour garantir que les apprenants soient libres

de choisir leurs propres mots, ce qui permet de s'approcher le plus possible des lacunes précises en langue des apprenants.

- **Systématique** : les corpus d'apprenants consistent en des collections systématiques de productions d'apprenants qui sont sélectionnées en fonction des critères externes (ex. : le niveau de langue, la langue maternelle des apprenants, le contexte d'apprentissage, le thème de production) (Nesselhauf, 2004). Ces critères de conception jouent donc un rôle important dans la mesure où ils déterminent la représentativité des corpus.

Les critères de conception qui influencent les productions des apprenants doivent être contrôlés afin de valoriser l'utilité des corpus d'apprenants. Il nous paraît important de prendre en compte ces critères dans la constitution du corpus et dans l'interprétation des résultats. Nous rejoignons Granger *et al.*, (2015) qui affirme que la collecte aléatoire de données hétérogènes ne peut pas être considérée comme un corpus d'apprenants. Cette affirmation rejoint la définition des corpus de Rastier que nous avons évoquée dans la section 2.1. On peut en déduire qu'il ne devrait pas y avoir de différences entre les corpus d'apprenants et les corpus de locuteurs natifs, les uns comme les autres étant composés de données structurées. Tenant compte des éléments mentionnés plus haut, nous adoptons la définition de Granger :

Computer learner corpora are electronic collections of authentic FL/SL textual data assembled according to explicit design criteria for a particular SLA/FLT purpose. They are encoded in a standardised and homogeneous way and documented as to their origin and provenance. (Granger, 2002, p. 7)

De cette définition, on retient que la constitution d'un corpus d'apprenants, prenant compte des aspects naturels et systématiques, doit d'abord répondre à un objectif pédagogique particulier dans le domaine de l'E/A des langues étrangères, ce qui conduit à l'apparition de différents types de corpus d'apprenants. Nous aborderons dans la section suivante les types de corpus d'apprenants, notamment dans le domaine de l'écrit académique afin de montrer l'intérêt de notre étude et de mieux situer notre corpus d'apprenants.

### 2.3.2.4. Le corpus Sup-Chinois : un corpus d'apprenants spécialisé sur l'écrit académique

Selon l'objectif pédagogique poursuivi, les corpus d'apprenants peuvent se subdiviser en plusieurs types : écrit, oral ou multimodal ; monolingue ou bilingue ; général ou spécialisé ; synchronique ou diachronique (Granger, 2002, p. 11). D'une manière générale, la plupart des corpus d'apprenants sont des écrits monolingues collectés pendant une certaine période, utilisés dans un objectif général d'apprentissage. Par ailleurs, l'anglais est la langue la plus représentée dans les corpus d'apprenants. Pour le français, les corpus d'apprenants restent moins nombreux, notamment du fait qu'il y a moins de demandes pour le FLE que pour l'EFL. Sur le site du *Centre for English Corpus Linguistics* (CECL)<sup>7</sup> est accessible un inventaire des corpus d'apprenants existants de nos jours.

Dans le domaine de l'écrit académique, les corpus d'apprenants sont largement utilisés par les linguistes et les enseignants dans l'identification des difficultés en matière d'utilisation du lexique propre à la démonstration scientifique (AWL, LST) d'une façon appropriée. À notre connaissance, le corpus *International Corpus of Learner English* (ICLE) (Granger *et al.*, 2015) est le plus grand corpus d'apprenants de l'anglais (comme langue étrangère) avec deux millions de mots issus de textes argumentatifs produits par des apprenants avancés ayant des langues maternelles diverses (allemand, français, chinois, etc.). L'équivalent oral est le *Louvain International Database of Spoken English Interlanguage* (LINDSEI) (De Cock, 2014). L'élaboration de ces deux corpus a été coordonnée par Granger et son équipe au *Centre for English Corpus Linguistics* (CECL) à l'Université Catholique de Louvain-la-Neuve.

Le corpus ICLE est utilisé dans un grand nombre de recherches<sup>8</sup> qui visent à identifier l'utilisation des structures grammaticales et des éléments lexicaux et discursifs dans les écrits académiques des apprenants. La variété des sous-corpus, composés des écrits produits par des apprenants de différentes langues maternelles

---

<sup>7</sup> Consulté le 20 mai 2016 à l'adresse suivante : <http://www.uclouvain.be/en-cecl-icworld.html>

<sup>8</sup> Liste des publications basées sur l'étude du corpus ICLE : <http://www.fltr.ucl.ac.be/FLTR/GERM/ETAN/CECL/learner%20corpus%20bibliography.html>

permet d'identifier les caractéristiques spécifiques à une ou plusieurs populations (Altenberg, 1997 ; Lorenz, 1998 ; Granger et Rayson, 1998).

En dehors du corpus ICLE, nous pouvons citer les projets suivants : *the Hong Kong University of Science & Technology learner corpus* (HKUST) (Milton & Chowdhury, 1994), *the Corpus of Academic Learner English* (CALE) (Callies & Zaytseva, 2013), *the Scientext English Learner Corpus* (Henderson, 2014). Dans la même lignée de ces travaux, on se focalise sur les besoins langagiers des apprenants dans leurs écrits académiques en français. Cependant, notre corpus est spécialisé en ce que nous nous focalisons sur le public des apprenants chinois, en vue de les aider à répondre aux attentes scripturales de l'institution.

Par ailleurs, on peut mentionner la constitution en cours d'un corpus d'apprenants au sein du laboratoire Lidilem. Il s'agit d'un corpus d'apprenants inscrit dans le domaine de l'écrit académique regroupant les productions écrites d'apprenants de diverses langues maternelles (le polonais, le vietnamien, le chinois, etc.).

Passons maintenant à la présentation du corpus d'étudiants natifs utilisé dans l'étude comparative.

### **2.3.3. Le corpus d'étudiants natifs : corpus Sup-Natifs**

Comme évoqué précédemment, afin de mener une comparaison avec le corpus Sup-Chinois, nous avons utilisé un corpus d'écrits d'étudiants natifs du français. Ce corpus, de 460 000 mots, est composé de 41 mémoires d'étudiants de master (1 et 2) en spécialité didactique du français. Il s'agit d'un sous-corpus du corpus Littéracie Avancée<sup>9</sup>, constitué pour analyser les compétences rédactionnelles en français langue maternelle à un niveau avancé et développer, par la suite, des ressources didactiques (Rinck & Jacques, 2015). Le choix de ce sous-corpus se justifie par le fait qu'il est comparable à notre corpus d'apprenants Sup-Chinois à différents égards. D'abord, la taille de ces deux corpus se rapproche, ce qui permet une comparaison quantitative. Ensuite, ils s'inscrivent dans le domaine de l'écrit académique, car ils sont constitués

---

<sup>9</sup> Corpus constitué par Rinck, F. Boch, F. & Jacques, M-P, consulté le 11/01/2016 à l'adresse suivante : <http://lidilem.u-grenoble3.fr/ressources/corpus-du-labo/article/corpus-litteracie-avancee>



en vue d'analyser les compétences rédactionnelles des étudiants en français. Enfin, ils partagent certaines marques linguistiques, par exemple des marqueurs d'organisation tels que *tout d'abord*, et des marqueurs énonciatifs comme *on* ou encore *nous*. Cependant, il faut souligner que certaines différences existent entre ces deux corpus. On revient sur cette question de comparabilité dans la section 2.3.5.

### 2.3.3.1. L'intérêt de l'utilisation du corpus d'étudiants natifs

Il est important de souligner que le corpus d'experts Scientext-SHS n'est pas comparable avec le corpus d'apprenants tant au niveau scientifique qu'au niveau des objectifs de communication. En effet, les articles de recherche rédigés par les experts de disciplines visent à diffuser les connaissances scientifiques tandis que les mémoires des étudiants (chinois et français) sont des écrits de recherche en formation requis pour valider leurs études de master.

Comme le soulignent Gilquin et Paquot (2008), par rapport aux corpus d'experts, les écrits des étudiants natifs fournissent des données mieux adaptées pour cibler les difficultés des apprenants non-natifs dans la mesure où ils permettent d'effectuer une comparaison d'une façon juste et égale. La question du type de modèle pour une comparaison reste entière :

On the one hand, it can be argued that in order to evaluate foreign learner writing by students justly, we need to use native-speaker writing that is also produced by students for comparison. On the other hand, it can also be argued that professional writing represents the norm that advanced foreign learner writers try to reach and their teachers try to promote. In this respect, a useful corpus for comparison is one which offers a collection of what Bazerman (1994: 131) calls 'expert performances'. (Ädel, 2006, p. 206-207, cité par Gilquin & Paquot, 2008, p. 45)

En effet, contrairement à certains chercheurs (Lorenz, 1999 ; Hyland & Milton, 1997) qui critiquent l'utilisation du corpus d'experts dans la recherche des corpus d'apprenants, nous considérons que les écrits des experts présentent une (variété de) norme, nous fournissant des ressources intéressantes dans les applications pédagogiques des constructions verbales pour développer les compétences discursives et argumentatives des apprenants. Les corpus d'étudiants natifs ne peuvent pas servir

de modèle parce que les étudiants natifs (universitaires), en tant que scripteurs novices en matière de rédaction scientifique, sont eux-mêmes confrontés aux difficultés liées à la maîtrise de l'écrit (par exemple, le mélange du registre) et à la construction des savoirs disciplinaires (par exemple, la difficulté à se positionner) (Rinck, 2011). Ainsi, dans une étude récente (Hatier & Yan, 2017) sur les constructions verbales de quatre verbes (*montrer, considérer, expliquer, décrire*), nous avons observé que les difficultés liées aux fonctions rhétoriques/discursives (absence de constructions verbales comme *s'expliquer par, considérons ces facteurs*) sont largement partagées par les étudiants natifs et non-natifs.

Afin de répondre à cette question, à l'instar de Granger et Paquot (2009), nous voulons intégrer dans notre analyse comparative à la fois le corpus d'experts et le corpus d'étudiants natifs parce qu'il nous semble que chacun d'eux présente un intérêt. D'une part, le corpus d'experts est considéré comme un modèle pour la description linguistique et pour les applications pédagogiques. Ici, nous empruntons la notion de modèle définie par Chambers (2010, p. 13) comme « une source d'exemples de l'usage des locuteurs natifs qui sont experts dans certains domaines, et non pas dans un sens prescriptif ». D'autre part, le corpus d'étudiants natifs constitue une référence, permettant de mettre en évidence les difficultés propres à l'apprenant chinois et celles qui sont aussi rencontrées par le scripteur français novice de l'écriture académique.

#### **2.3.4. Les choix des méthodes d'analyse pour le diagnostic des besoins langagiers**

Nous allons maintenant nous arrêter sur les méthodes d'exploitation de corpus d'apprenants. Compte tenu de nos objectifs liés à l'analyse linguistique des besoins langagiers, nous avons adopté deux méthodes d'analyse dans l'analyse des corpus d'apprenants : l'analyse des erreurs et l'analyse des résultats de comparaisons de corpus.

La première méthode se rapporte à l'analyse des erreurs assistée ou non par ordinateur. D'une part, les productions des apprenants recueillies sont traitées directement dans un concordancier et les erreurs contenues peuvent être relevées

manuellement d'après une typologie d'erreurs établie (Anctil, 2005 ; Luste-Chaâ, 2010) ; d'autre part, lorsqu'il s'agit de l'analyse des erreurs assistée par ordinateur, les corpus doivent être annotés sur les erreurs (tels que FRIDA, HKUST et ICLE), en suivant des typologies d'erreurs existantes. Le travail d'étiquetage permet ensuite de détecter quelques erreurs automatiquement (L'haire & Vandeventer-Faltin, 2003 ; Chuang & Nesi, 2006). L'étude de Granger (2007) sur le corpus FRIDA, vise à construire un système d'étiquetage et à mettre en place un outil de détection et de correction automatique des erreurs produites par les apprenants en FLE. Notons que dans les deux cas d'analyse des erreurs, la correction manuelle des corpus d'apprenants et la construction d'une typologie d'erreurs sont incontournables.

Notre corpus d'apprenants ne comporte pas d'annotation des erreurs. Nous choisissons donc d'interroger directement dans un concordancier notre corpus d'apprenants pourvu d'annotations morphosyntaxiques. En effet, comme évoqué précédemment, nous n'envisageons pas de relever toutes les erreurs et notre analyse des erreurs est centrée sur les verbes du LST. Par ailleurs, on ne dispose pas de grille sur les constructions à l'heure actuelle, si bien que le travail d'annotation paraît prématuré. L'analyse des erreurs relevées dans notre corpus d'apprenants sera abordée dans le chapitre V en fonction des objectifs visés présentés précédemment.

La deuxième méthode nommée Contrastive Interlanguage Analysis (CIA) (Granger, 2002) consiste à effectuer des comparaisons quantitative et qualitative entre les productions des locuteurs natifs et non-natifs ou celles entre les différents groupes de locuteurs non-natifs. Cette approche contrastive nous intéresse plus particulièrement puisqu'elle permet d'effectuer un repérage des cas de suremploi ou de sous-emploi d'éléments langagiers spécifiques dans les productions des apprenants (par rapport aux scripteurs natifs). Comme le soulignent Granger et Paquot (2009), les difficultés des étudiants non-natifs méritent d'être prises en compte et traitées spécifiquement. Il en est certes de même pour les natifs, mais les erreurs portent sur des éléments différents. Ces aspects seront développés dans le chapitre V où nous présenterons les travaux portant sur la comparaison entre les écrits des natifs/non-natifs et examinerons l'emploi des verbes et des constructions verbales du LST chez les étudiants natifs/non-natifs en nous basant sur le corpus Sup-Chinois et le corpus Sup-Natifs, et en les comparant au corpus d'experts Scientext-SHS.

### 2.3.5. Le tableau récapitulatif de comparabilité

Nous prenons pour corpus d'analyse différents ensembles de textes se rapportant aux genres de l'écrit scientifique et de l'écrit académique, le genre étant défini ici dans un sens large, à la suite de Biber (1993), comme une catégorie de textes de même fonction communicative.

Les deux corpus de mémoires d'étudiants chinois et français, même s'ils portent la même dénomination, n'illustrent pas des pratiques identiques, car les mémoires d'étudiants chinois ne suivent pas une démarche expérimentale. La méthodologie y est absente et l'analyse scientifique consiste notamment à mener une discussion culturelle et linguistique. De ce fait, il nous semble que les mémoires d'étudiants chinois se rapprochent du genre de la dissertation. Par contre, ces deux corpus relevant de l'écriture d'apprentissage peuvent être regroupés dans un même sous-genre dit d'« écrits de recherche en formation » (Reuter, 2004). Comme le soulignent Cadet et Rinck (2014, p. 4), ces mémoires produits dans le contexte académique présentent une logique interne et externe (écritures codifiées) en posant « la question de leurs relations avec “les écrits pour penser” et des relations entre savoirs universitaires, savoirs scolaires et vécus ».

Le corpus d'experts appartient, quant à lui, au sous-genre de l'article de recherche. Comme nous l'avons déjà indiqué dans la section 2.3.1., l'article de recherche en tant que modèle (ou support de l'E/A) retenu se caractérise par sa relative brièveté et sa représentativité dans la construction et dans l'écriture, permettant des descriptions linguistiques et applications pédagogiques (Pollet, 2014, p. 11).

Rappelons que, pour l'ensemble de ces genres d'écrits, nous nous intéressons au niveau lexical transdisciplinaire. Sur cette dimension linguistique, les articles de recherche ont beaucoup de points communs avec les écrits de recherche en formation, ce qui rend l'analyse comparative faisable. Néanmoins, nous sommes consciente que les trois corpus d'analyse ne sont pas tout à fait comparables en matière de genre. Lors de l'étape d'analyse contrastive, nous tiendrons compte de ces différences de genre dans les corpus.

Comme nous l'avons déjà expliqué, notre objectif en comparant ces trois corpus est d'aboutir au diagnostic des difficultés liées aux verbes et aux constructions verbales du LST pour les apprenants chinois, en prenant en considération les éléments lexicaux et syntaxiques à enseigner dans les cours du FLE.

Ci-dessous un tableau récapitulatif illustre les principaux critères de conception pris en compte dans les trois corpus :

Corpus	Taille	Annotation	Genre	Public	Discipline	Type de texte
Scientext-SHS	5 000 000 mots	annotation morphosyntaxique, annotation syntaxique en dépendances	article de recherche	experts francophones	10 disciplines sciences humaines et sociales	argumentatif
Sup-Chinois	600 000 mots	annotation morphosyntaxique	écrits de recherche en formation (mémoires d'étudiants de 5 <sup>e</sup> année)	apprenants chinois	FLE	argumentatif
Sup-Natifs	460 000 mots	annotation morphosyntaxique	écrits de recherche en formation (mémoires français M1, M2)	étudiants français	didactique du FLE	argumentatif

**Tableau 2.2 : Récapitulatif de la comparabilité de trois corpus d'analyse**

Comme indiqué dans le tableau, le corpus Sup-Natifs est plus proche du corpus Sup-Chinois, notamment aux niveaux de la taille, du genre et du type de texte. Il servira donc de référence pour la comparaison. Par ailleurs, le corpus Scientext-SHS, avec une taille beaucoup plus importante, constitue un modèle d'experts que les scripteurs novices souhaitent atteindre, car il fournit les éléments linguistiques qui nous intéressent en termes de fréquences.

Les différences entre les trois corpus, notamment celles liées aux disciplines seront prises en compte dans l'interprétation des données. Nous expliquons dans la section suivante la méthode et les outils d'exploitation pour les trois corpus d'analyse.

## 2.4. Les outils d'exploitation

Dans cette partie, nous présentons trois outils d'exploitation utilisés dans notre

analyse des corpus : ScienQuest, Lexicoscope et TXM. Les deux premiers s'appliquent au corpus Scientext-SHS tandis que le dernier permet d'interroger les deux corpus de mémoires d'étudiants.

### 2.4.1. ScienQuest

Un sous-ensemble du corpus Scientext-SHS (environ 300 articles) est consultable en intranet avec l'outil ScienQuest<sup>10</sup> (Falaise *et al.*, 2011a ; Tutin & Falaise, 2013). Initialement développé pour les sous-corpus<sup>11</sup> du projet Scientext, il est « conçu pour effectuer sans connaissances techniques préalables des recherches sur les parties textuelles, les parties du discours et les fonctions syntaxiques » (Falaise *et al.*, 2011a, p. 103).

L'outil ScienQuest n'intègre qu'une partie de notre corpus d'analyse, soit environ 3 millions de mots relevant d'environ 300 articles scientifiques de 10 disciplines de SHS. Ceci n'est pas gênant puisque nous utilisons ScienQuest notamment pour une vérification manuelle. Le travail de l'extraction des verbes et des cadres de sous-catégorisation se fait d'une façon semi-automatique à l'aide des techniques de TAL (*cf.* Chapitre III). L'outil ScienQuest nous permet de chercher les constructions syntaxiques récurrentes liées à un verbe du LST ou les collocations au niveau des arguments sélectionnés par le verbe. Nous présentons ci-dessous l'exploitation du corpus à l'aide de l'outil ScienQuest qui se réalise en trois étapes : choix d'un corpus, formulation de requêtes, traitement des résultats.

#### 2.4.1.1. Méthodes d'exploration du corpus

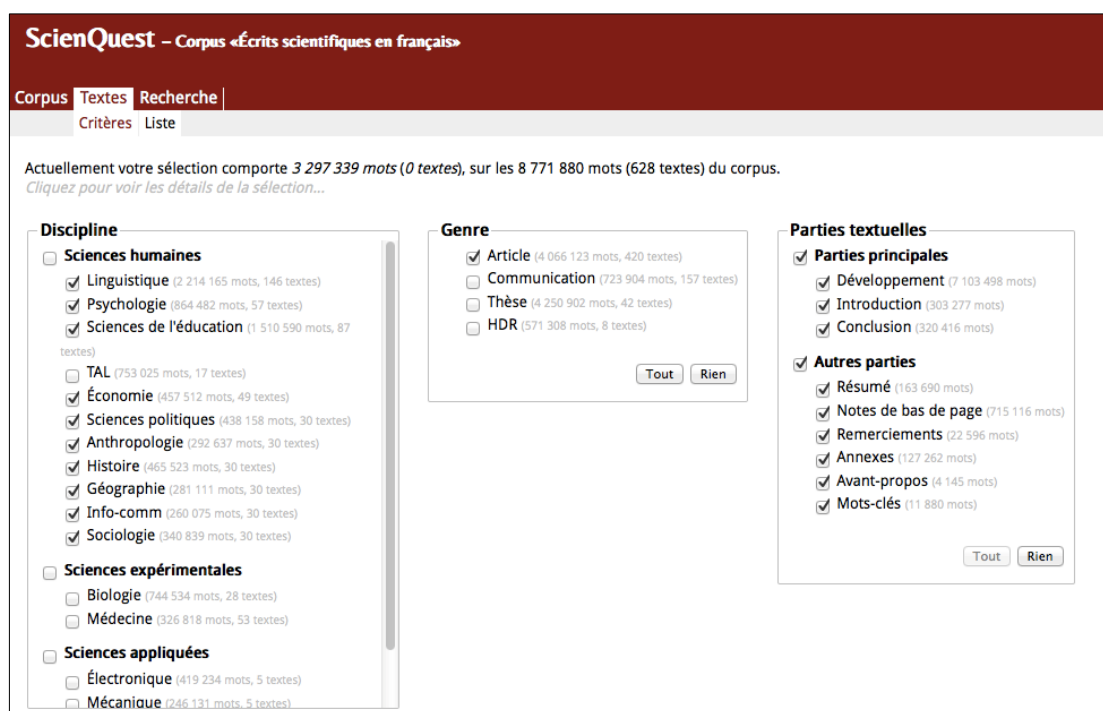
Sur l'interface ScienQuest, pour faire des recherches de verbes ou de constructions syntaxiques, la première étape consiste à sélectionner le corpus d'analyse. La figure 2-2 montre que l'interface propose trois possibilités de sélections par disciplines, genres et parties textuelles. Dans notre cas, nous nous intéressons aux

---

<sup>10</sup> <http://corpora.aiakide.net/scientext18>

<sup>11</sup> La plateforme comporte un ensemble de sous-corpus en ligne : un corpus français d'écrits scientifiques, un corpus anglais d'écrits scientifiques, un corpus anglais d'apprenants et un corpus expérimental de commentaires évaluatifs. En outre, un corpus français d'apprenants est en cours de construction.

patrons verbaux dans les articles de recherche relevant de 10 disciplines SHS, toutes les parties textuelles confondues.



**Figure 2-2 : Sélection d'un sous-corpus dans ScienQuest pour le corpus français de Scientext : articles de recherche dans 10 disciplines SHS**

Après avoir délimité le corpus d'analyse, nous pouvons accéder aux requêtes par trois modes de recherche fondés sur le même langage du moteur de recherche ConcQuest<sup>12</sup> développé par Kraif et Falaise (Falaise *et al.*, 2011b) : une recherche sémantique qui fournit un accès onomasiologique, renvoyant à des fonctions sémantiques spécifiques préétablies (ex. : formulation d'une hypothèse), une recherche libre guidée qui permet de définir la requête aux niveaux morphosyntaxiques et syntaxiques (ex. : le verbe *montrer* accepte quel type de noms comme objet direct), et une recherche avancée qui utilise des expressions régulières et les dépendances syntaxiques dans une requête complexe (ex. : création de requêtes complexes avec des relations syntaxiques variées).

Nous adoptons dans notre étude le mode de recherche libre guidée en raison de sa simplicité d'accès (notamment par rapport à la recherche avancée qui est davantage réservée aux linguistes familiarisés avec le TAL) qui permet de sélectionner des

<sup>12</sup> Un logiciel d'extraction de concordances permettant de rechercher des expressions complexes.

formes, des lemmes, des catégories et des relations syntaxiques simplement. En outre, il est possible de limiter le nombre d'occurrences à 100 mots ou plus, et paramétrer les contextes gauche et droit selon les besoins. La figure 2-3 illustre la recherche pour le verbe *observer* ayant comme sujet les pronoms *nous/on/je* et comme objet direct un nom.

The screenshot shows the ScienQuest search interface. At the top, there are tabs for 'Corpus', 'Textes', 'Recherche', and 'Résultats'. Below these, there are search criteria for three words: Mot 1 (Forme, Lemme: nous/on/je, Catégorie), Mot 2 (Forme, Lemme: observer, Catégorie), and Mot 3 (Forme, Lemme, Catégorie: Nom, Traits). A 'Relations syntaxiques' section shows 'Mot 1' as 'sujet de (SUI)' and 'Mot 3' as 'objet direct de (OBJ)'. A search button and '137 occurrences' are visible. Below is a table with 8 rows of results, each showing a context, the occurrence, and the text.

#	Contexte gauche	Occurrences	Contexte droit	Texte
1	sociologues qualifiant d' effets pervers ". De fait , lorsqu'	on observe la réception	de l' anthropologie américaine en France , on constate que	#9 - Article - Anthropologie - Développement
2	bénéficier d' un regard extérieur a facilité les observations .	Je me suis attachée à observer les échanges	entre professionnels lors de ce que Barrett [ 1996 ] appelle	#14 - Article - Anthropologie - Développement
3	autres professionnels et de contrôler le recrutement des enfants .	On observe un fort pourcentage	d' enfants issus de milieux modestes ou défavorisés , pour	#14 - Article - Anthropologie - Développement
4	, je voyais mon père jouer de la tabla ,	j' observais ses mains	quand il jouait et j' étais vraiment passionné par l'	#34 - Article - Anthropologie - Développement
5	dérivée par rapport à ln ( ) ". Par ailleurs , si	on observe en détail la courbe	tracée sur le graphique bilogarithmique , il apparaît très clairement	#49 - Article - Géographie - Développement
6	bougé le petit doigt ". Ainsi , si l'	on observe la dispersion	des voix , dans chacun des deux principaux partis ,	#50 - Article - Géographie - Développement
7	régions mondiales ( 1975-2005 ) Entre 1975 et 2005 ,	on observe des recompositions	majeures des échanges de produits manufacturés selon la dimension centre	#51 - Article - Géographie - Développement
8	région ( - 0,13 ) . En asie pacifique ,	on observe également un transfert	vers la périphérie , mais sans pour autant affaiblir la position	#51 - Article - Géographie - Développement

**Figure 2-3 : Requête pour le verbe *observer* ayant comme sujet les pronoms *nous/on/je* et objet les noms**

En outre, avec un simple clic, le concordancier nous permet d'accéder au contexte élargi et aux dépendances syntaxiques de la phrase affichées sous forme graphique.

### 2.4.1.2. Traitement des résultats

Dans ScienQuest, une fois les requêtes terminées, les concordances affichées dans un format classique KWIC (Key Word In Context) peuvent être filtrées, sauvegardées, puis exportées dans un tableur (format HTML, CSV ou XLSX) pour effectuer une analyse notamment sémantique des patrons verbaux. Les statistiques sur les occurrences trouvées peuvent être obtenues en fonction de différents critères : forme, lemme, catégorie, traits morphosyntaxiques, partie textuelle, genre et discipline. Par exemple, la figure 2-4 indique le nombre d'occurrences et le



pourcentage des constructions verbales (affichées en lemme) pour le verbe *observer* ayant comme sujet les pronoms *nous/on/je* et un objet direct nominal dans un ordre décroissant. On peut constater que les constructions *on observe différence/tendance/effet*, *nous observons différence/corrélation* sont les plus fréquentes réparties dans les 10 disciplines SHS. Il est également possible de voir la répartition par discipline (voir la figure 2-5).

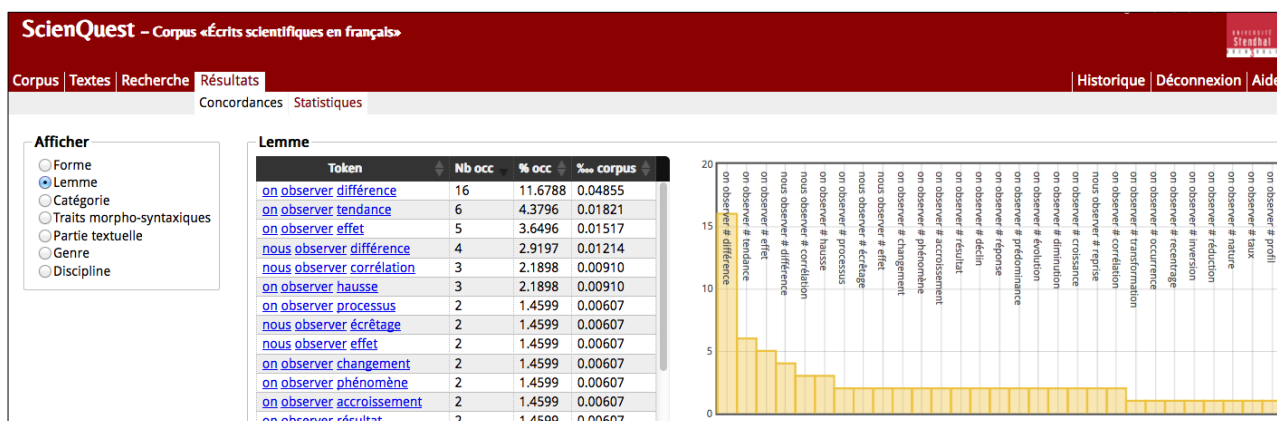


Figure 2-4 : Visualisation en lemme des constructions verbales *nous/on/je* + *observer* + noms

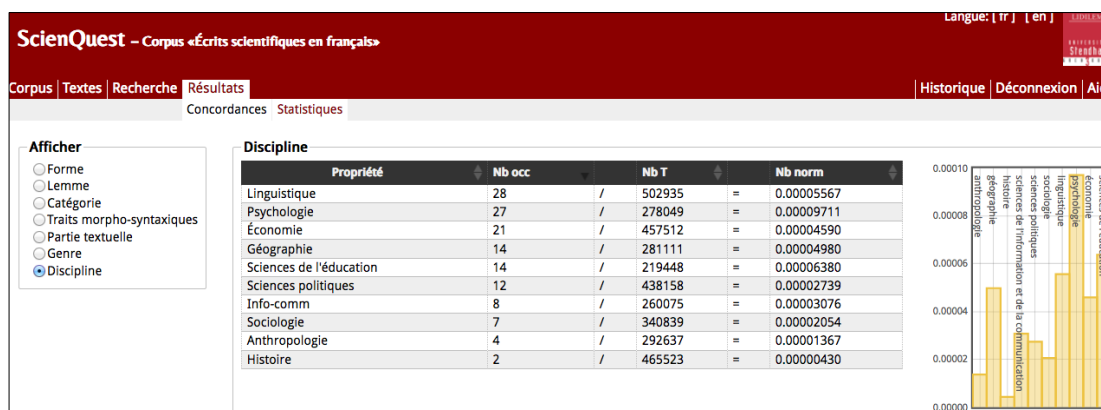


Figure 2-5 : Visualisation de répartition par disciplines

En dehors de l'outil ScienQuest, nous avons recours à l'outil Lexicoscope pour compléter notre recherche sur le corpus Scientext-SHS. En effet, cet outil a l'avantage d'extraire les co-occurents du mot, ce qui nous permet d'examiner les arguments nominaux du verbe du LST.

## 2.4.2. Le Lexicoscope


L'ensemble du corpus d'analyse Scientext-SHS a été intégré dans la plateforme Lexicoscope<sup>13</sup>, un outil en ligne développé au Lidilem (Kraif et Diwersy, 2012). Le Lexicoscope permet d'extraire à la fois des concordances et des lexicogrammes, c'est-à-dire des tables de co-occurrences, ce qui est pratique pour étudier le profil combinatoire des unités phraséologiques. Dans la section suivante, nous montrons quelques fonctionnalités que nous avons utilisées dans l'analyse des verbes du LST.

### 2.4.2.1. Recherche des co-occurents

D'abord, la requête libre permet de repérer facilement tous les co-occurents du mot donné. Si nous souhaitons connaître les co-occurents relevant d'une certaine catégorie grammaticale (préposition, adverbe, nom, etc.) ou d'une position syntaxique (sujet, objet, etc.), nous pouvons utiliser la requête multi-pivots. Par exemple dans la figure 2-6, la case relation permet de préciser que l'adverbe est dépendant du verbe. Les co-occurents d'adverbes les plus fréquents du verbe *analyser* sont *ici*, *notamment*, *ensuite*, *précisément* (voir la figure 2-7).

Figure 2-6 : Requête pour les co-occurents d'adverbes du verbe *analyser*

<sup>13</sup> Cet outil permet d'avoir, pour un mot donné, l'ensemble de ses co-occurents les plus significatifs, il est accessible en ligne : <http://dip01.u-grenoble3.fr/~kraifo/lexicoscope/>

**Lexicogrammes : Tableau de contingence et mesures d'association** 

Lexicogramme **Graphiques**

Show  entries Search:

I1	I2	f.deprels	f	f1	f2	N	f.disp	am.log.likelihood	r.log.likelihood
analyser_VERB	ici_ADV	U3_ADVVMOD	24	215	2183	126337	6	51,1879	1
analyser_VERB	notamment_ADV	U3_ADVVMOD	11	215	1710	126337	7	13,4254	3
analyser_VERB	ensuite_ADV	U3_ADVVMOD	9	215	797	126337	6	19,1275	2
analyser_VERB	précisément_ADV	U3_ADVVMOD	6	215	597	126337	4	11,5019	4

Showing 1 to 4 of 4 entries ◀ ▶

**Figure 2-7 : Visualisation des co-occurents d'adverbes du verbe *analyser***

Ces données se révèlent utiles pour interpréter les valeurs sémantiques du verbe.

#### 2.4.2.2. Recherche des constructions syntaxiques

La requête avancée nous intéresse plus particulièrement puisqu'elle permet d'une part de relever les co-occurents liés à un mot donné et d'autre part de chercher les structures syntaxiques avec les co-occurents de mots qui apparaissent dans certaines positions syntaxiques de la phrase. La figure 2-8 illustre la requête pour le verbe *analyser* ayant comme sujet les pronoms et un objet direct nominal. Notons par ailleurs que les relations de dépendance pour chaque requête sont illustrées dans une figure en bas de l'écran. Les lignes des concordances nous permettent de constater que les sujets nominaux renvoient souvent aux pronoms *on*, *nous* et *je* qui représentent l'auteur (ou les auteurs) et que les objets nominaux relèvent plutôt de quelque chose d'abstrait comme *effet*, *évolution*, *cas*, etc (voir la figure 2-9). De ce fait, la requête avancée nous permet d'analyser le comportement lexico-syntaxique du verbe dans un contexte précis.

Sélection du Corpus Requête Paramètres Sessions sauvegardées Guide

### Concordances et profils combinatoires (cooccurrences)

Requête libre Requête avancée Requête multi-pivots

Expressions  
`<l=analyser,c=VERB,#1>&&<c=NOUN,#2>&&<c=PRON,#3>::  
 (.*OBJ.* ,1,2)(.*SUBJ.* ,1,3)`

Relations :  
 .\*  
 ~rel signifie que le pivot est en position de dépendant

Concordances Cooccurrences

Taper une expression utilisant le langage de requête avancé. P.ex. "`<l=considération,c=N,#1>`" ou "`<l=prendre,c=V,#1>&&  
 <l=en,c=PREP,#2>&&<l=considération,c=N,#3>::(adv,1,3) (pm,3,2)`"

Figure 2-8 : Requête pour le verbe *analyser* ayant comme sujet les pronoms et objet les noms

Nombre total d'occurrences 90  
 Dispersion 10

### Requêtes

- Requête : `<l=analyser,c=VERB,#1>&&<c=NOUN,#2>&&<c=PRON,#3>:: (.*OBJ.* ,1,2)(.*SUBJ.* ,1,3)`

Show 10 entries Search:

Identifiant	Contexte gauche	Pivot	Contexte droit
XML_anthropologie.xml-s1009	Elle	analyse	leurs effets déstabilisateurs dans la définition de soi comme homme et comme femme à la pêche et les tentatives de recomposition des repères identificateurs aujourd'hui à l' Suivre .
XML_geo.xml-s10594	Si on	analyse	l' évolution de la population depuis 1971 ( tab .
XML_sociologie.xml-s11010	J'	analyse	ensuite les discours tenus par ces jeunes qui racontent en quoi cette bifurcation représente pour eux un « carrefour , un moment où ils ont dû faire des choix » , pour identifier les ingrédients mobilisés dans leur prise de décision .
XML_anthropologie.xml-s11074	J'	analyserai	un cas particulier , celui d' une jeune thérapeute - possédée .
XML_ling.xml-s11253	Des huit interventions que j'	analyse	, j' ai constaté que ces constructions peuvent se diviser en deux catégories selon le rôle qu' elles jouent au sein de ces interventions .
XML_ling.xml-s11493	8 Etant donné que l' échange qu' on	analyse	est monologal et pas dialogal , il est normal de ne pas trouver ce genre d' implicitation .
XML_ling.xml-s11496	11 Roulet ne fait pas cette distinction car il a toujours	analysé	des échanges entre deux interlocuteurs où les notions de destinataire direct et indirect se superposent .
XML_scienceseducation.xml-s11500	Nous cherchons par là à respecter le caractère fondamentalement analogique des données que nous	analysons	.

Figure 2-9 : Lignes de concordances pour le verbe *analyser* ayant comme sujet les pronoms et objet les noms

Rappelons que sur le plan linguistique, le corpus Scientext-SHS est utilisé pour

observer les constructions syntaxiques associées au verbe du LST ainsi que les collocations au niveau des arguments sélectionnées par le verbe. Nous envisageons d'extraire ces informations de manière semi-automatique avec des techniques de TAL, puisque le travail d'extraction automatique présente certaines limites (*cf.* Chapitre III) et le recours aux outils ScienQuest et Lexicoscope nous permet de vérifier et d'affiner les données extraites, facilitant par conséquent l'interprétation sémantique des constructions verbales.

Reste à présenter l'outil TXM qui nous a servi pour l'analyse des corpus Sup-Chinois et Sup-Natifs.

### 2.4.3. TXM

#### 2.4.3.1. Interrogation des corpus avec le logiciel TXM

Les deux corpus d'étudiants ont été annotés au plan textuel par deux linguistes qualifiés<sup>14</sup> sous le format TEI Lite (Text Encoding Initiative), avec le langage de balisage XML. En outre, un étiquetage morphosyntaxique a été réalisé<sup>15</sup> avec l'étiqueteur MElt<sup>16</sup> pour rendre possible l'extraction automatique d'informations grammaticales.

Afin d'examiner l'emploi des constructions verbales du LST dans ces deux corpus, nous avons procédé à une analyse manuelle à l'aide du logiciel Textométrie (TXM)<sup>17</sup>, conçu par Heiden et ses collègues (Heiden, Magué, & Pincemin, 2010). Ce logiciel a l'avantage de proposer des requêtes en Corpus Query Language (CQL) et par conséquent de créer des concordances pour des schémas lexico-grammaticaux complexes (Pecman, 2012, p. 122). Nous donnons ci-dessous deux exemples de requêtes.

---

<sup>14</sup> Le corpus d'apprenants chinois est annoté par Clarisse Bayol, étudiante de Master en sciences du langage à l'Université Stendhal, celle-ci est encadrée par Thi Thu Hoai Tran ; le corpus d'étudiants français est annoté par Marie-Paule Jacques dans le cadre du projet « Littéracie avancée ».

<sup>15</sup> Nous remercions Marie-Paule Jaques pour son travail d'étiquetage morpho-syntaxique dans le corpus Sup-Chinois.

<sup>16</sup> Un étiqueteur statistique librement disponible en ligne : <http://gforge.inria.fr/projects/lingwb/>

<sup>17</sup> <http://textometrie.ens-lyon.fr/>

- (1) [pos="NC"[]]{0,4}[lemma="décrire" & (pos="VPP")]

*modèle décrit précédemment, X est décrit comme complexe, les activités sont décrites ci-dessous, etc.*

- (2) [lemma="comme"[]]{0,5}[lemma="montrer|indiquer|souligner|expliquer"]

*comme je l'ai indiqué plus haut, comme nous l'explique X, comme on l'a souligné dans le texte, etc.*

Comme on peut le constater, il est possible de faire une requête sur une propriété du mot (catégorie grammaticale), une combinaison d'informations (croisement d'une catégorie et d'un trait morphologique) ainsi que des séquences d'unités lexicales (insertion ou distance). La première requête permet d'extraire les occurrences du verbe *décrire* au participe passé et précédé d'un nom commun avec une distance de zéro à quatre. La deuxième requête permet d'extraire les occurrences des constructions en *comme* suivis des verbes *montrer*, *indiquer*, *souligner* et *expliquer*. De cette manière, nous pouvons donc vérifier si les formes verbales ou les constructions verbales récurrentes dans les écrits scientifiques se rencontrent dans les deux corpus d'étudiants. La figure 2-10 illustre la deuxième requête dans le corpus Sup-Natifs :

text_id	Contexte gauche	Pivot	Contexte droit
linguistique_1_tag	il s'agit là d'une implicature,	comme l'indique	le caractère annulable de l'information sous-entendue : « À cette époque
littérature_11_tag	espace poétique sous la plume de Baudelaire est	comme un gouffre irrésistible qui montre	aux lecteurs son immensité. lorsque la poésie de Verlaine s'exprime
littérature_11_tag	rare. Huymans connaît cependant déjà Corbière,	comme l'indique	sa lettre d'octobre 1882 à Mallarmé. Son commentaire atteste une
littérature_11_tag	poètes. 3.3.2 Une poésie obscure et concentrée	Comme l'indique	clairement la lettre que Huymans envoie à Mallarmé en octobre 1882,
littérature_12_tag	le temps au roman est diminué extrêmement.	Comme ce que Ephraïm Gotthold Lessing souligne	, l'espace appartient au domaine de peinture. Simon utilise le
littérature_2_tag	plutôt floued laquelle Bakhtin définit le concept,	comme le souligne	Bart Keunen. Gabriel Zoran, a publié Towards a Theory of
littérature_2_tag	le lieu de l'expérimentation. Dans le roman	comme recherche, il souligne	: « Alors que le récit véridique a toujours l'appui,
littérature_2_tag	sur une surface mais pas un fil.	Comme Butor a souligné	dans son article Recherches sur la technique du roman, la narration
littérature_5_tag	que son entourage, origine des stimulations.	Comme explique	l'auteur : « l'élément psychologique, comme l'élément pictural
socioculturel_5_tag	de la mer se sont font l'objet	comme des zones économiques spéciales indiquées	particulièrement par Dengxiaoping au but d'attirer l'investissement étranger pour le
traduction_2_tag	sauver. Il cueille une branche de saule	comme cadeau. Il doit s'expliquer	et se défendre. Si l'on ne met pas ce verbe
traduction_4_tag	l'objectivité de la part du traducteur,	comme on l'indique	: Son choix du genre de l'oeuvre et de la stratégie
traduction_4_tag	se guélier », qui est souvent utilisé	comme adjectif pour indiquer	quelque chose de favorable et encourageant. A travers cet exemple,
traduction_4_tag	il juge utile à une compréhension parfaite,	Comme ce que Jacqueline Henry le souligne	, " de telles notes sont parfois permises, et parfois même
traduction_7_tag	langue et la culture de d'arrivée.	Comme nous avons indiqué	dans l'introduction précédente, cela détermine donc que la traduction se

**Figure 2-10 : Requête des constructions en *comme* avec les verbes *montrer*, *indiquer*, *souligner* et *expliquer* dans le corpus Sup-Natifs**

Les occurrences peuvent être affichées en associant leur forme et leur propriété morphosyntaxique. Il y a 192 occurrences et la colonne « pivot » recense les mots recherchés. Par défaut, les concordances sont triées par ordre alphabétique de la

colonne « pivot ». Notons en outre que les concordances peuvent être exportées au format CSV. Pour résumer, l'exploration de deux corpus d'étudiants associe l'analyse manuelle et automatique et permet ainsi d'étudier l'emploi des verbes et des constructions verbales du LST dans les écrits des scripteurs novices par rapport aux experts.

## 2.5. Bilan

Dans ce chapitre, nous avons présenté les données collectées dans notre étude. Nous nous sommes positionnée d'abord par rapport à la vision sur la linguistique de corpus qui constitue de nos jours la méthodologie privilégiée de recherche linguistique. Néanmoins, on ne peut pas se passer des théories linguistiques existantes et négliger l'importance des interprétations des données. Dans le domaine de l'écrit scientifique et académique, il a été démontré que les corpus sont utiles non seulement pour comprendre le fonctionnement de la langue du genre, mais aussi pour favoriser son apprentissage.

Nous avons introduit successivement les corpus qui jouent chacun un rôle dans notre analyse linguistique avec un objectif didactique précis tel que l'aide à la rédaction académique. Nous nous sommes concentrée plus particulièrement sur la présentation du corpus d'apprenant Sup-Chinois qui permet d'effectuer une étude de diagnostic des besoins langagiers. Comme nous l'avons montré, dans le domaine du FLE, il existe peu de corpus d'apprenants inscrits dans un genre précis comme l'écrit académique, et ayant pour objectif le diagnostic des besoins langagiers. Il s'agit donc d'un aspect qui montre l'intérêt de notre étude. Nous avons ensuite abordé la question de la comparabilité des corpus dans le cadre de notre étude. Certes, il existe différents critères de comparabilité, que ce soit qualitatifs ou quantitatifs, mais ils ne peuvent pas tous être pris en compte. En ce qui concerne notre étude, la comparabilité des trois corpus est basée à la fois sur certaines caractéristiques communes aux textes et sur les objectifs applicables des corpus. Enfin, nous avons illustré l'utilisation de plusieurs outils d'exploitation (ScienQuest, Lexicoscope et TXM) qui facilitent notre analyse linguistique sur corpus.

---

## Chapitre III.

# Les patrons lexico-syntaxiques des verbes du LST : de l'extraction à la modélisation

---

3.1. EXTRACTION DES VERBES DU LST ET DES CADRES DE SOUS-CATEGORISATION .....	93
3.1.1. Extraction des verbes du LST .....	93
3.1.2. Extraction des cadres de sous-catégorisation par TAL .....	94
3.2. TRAITEMENT SEMANTIQUE DES VERBES DU LST .....	97
3.2.1. Repérage des acceptions avec la ressource LVF .....	98
3.2.2. Classification sémantique .....	104
3.3. UTILISATION DU MODELE CPA DANS L'ECRIT SCIENTIFIQUE .....	108
3.3.1. Attribution des types/rôles sémantiques.....	109
3.3.2. Vers des patrons syntaxiques spécifiques au genre .....	114
3.3.3. Représentation des patrons par entrée onomasiologique.....	120
3.3.4. Récapitulatif des points communs et des différences.....	122
3.4. BILAN .....	123

---

Ce chapitre, consacré aux méthodes d'extraction et d'analyse linguistique des verbes du LST, est composé de trois parties. Premièrement, nous présentons le travail d'extraction des verbes du LST et des cadres de sous-catégorisation qui servira de base de données pour la modélisation des patrons verbaux. Deuxièmement, nous expliquons les traitements sémantiques des verbes du LST en nous basant sur la ressource lexicographique *Les verbes du français* (LVF). L'objectif est d'établir une classification des acceptions des verbes du LST. Enfin, nous cherchons à démontrer de quelle façon le modèle CPA peut s'appliquer à notre étude qui vise un objectif plus spécifique, celui d'analyser les verbes dans l'écrit scientifique en français. La question que nous posons concerne les spécificités à prendre en compte au cours de la modélisation des patrons verbaux dans l'écrit scientifique.



### 3.1. Extraction des verbes du LST et des cadres de sous-catégorisation

Rappelons que, sur le plan linguistique, nous envisageons d'observer le comportement lexico-syntaxique des verbes du LST dans le corpus Scientext-SHS et de l'analyser avec le modèle des patrons CPA. L'analyse consiste donc à dresser un inventaire des principaux patrons des verbes du LST. Nous nous démarquons cependant de la méthode d'analyse de Hanks qui analyse des échantillons de concordances manuellement<sup>1</sup>, en adoptant un repérage semi-automatique des patrons. Cette procédure se réalise en trois étapes principales :

- (1) Extraction des verbes du LST et des cadres de sous-catégorisation
- (2) Identification des sens du verbe à partir des cadres extraits et des lignes de concordances du corpus Scientext-SHS, en faisant des correspondances avec la ressource LVF
- (3) Créations des patrons en adaptant le modèle CPA à notre objectif spécifique

Nous présentons ci-dessous l'étape d'extraction.

#### 3.1.1. Extraction des verbes du LST

Le travail d'extraction du lexique scientifique transdisciplinaire a été réalisé par Hatier (2016) dans le cadre de sa thèse intitulée « Extraction et catégorisation de lexiques scientifiques transdisciplinaires d'articles scientifiques de sciences humaines » au sein du projet ANR TermITH. Des listes de mots simples (noms, verbes, adjectifs et adverbes) sont constituées d'une façon semi-automatique en combinant les techniques de traitement automatique des langues (TAL) et la vérification manuelle par les linguistes.

Le processus d'extraction du lexique est basé sur le corpus Scientext-SHS de 5

---

<sup>1</sup> Dans le PDEV, la construction des patrons lexico-syntaxiques d'un verbe se réalise à partir de 250 lignes de concordances aléatoires, un numéro identique du patron est attribué à chaque ligne de concordance partageant le même sens.

millions de mots et sur un corpus de contraste de 150 millions de mots (écrits journalistiques, fiction, oral transcrit, sous-titres), répondant à des critères de transdisciplinarité et de spécificité. Ainsi, un mot peut être considéré comme appartenant au LST s'il répond aux critères suivants :

1. il est surreprésenté dans le corpus Scientext-SHS par rapport au corpus de contraste, et ceci dans au moins 3 disciplines ;
2. il est présent dans un minimum de 50 tranches sur 100 dans le corpus d'analyse.

La liste des mots ainsi obtenue est ensuite soumise à la validation finale par des experts linguistes en s'appuyant sur des exemples phrastiques issus de différentes disciplines et sur les associations lexico-syntaxiques les plus significatives extraites des analyses en dépendances (Hatier, et *al.*, 2016). À l'issue de la validation manuelle, on obtient une ressource du LST composée de 1297 mots pleins (noms, verbes, adjectifs et adverbes) dont 342 verbes simples. La liste des verbes du LST est détaillée dans l'Annexe 1. Par ailleurs, cette liste du LST est enrichie par une classification sémantique en classes et sous-classes transcatégorielles (noms, verbes, adjectifs et adverbes)<sup>2</sup> à travers un travail manuel d'identification des sens présents dans le corpus Scientext-SHS, ce qui permet de passer d'une liste de lemmes du LST à une liste ayant pour entrées des acceptions. Nous reviendrons sur ce point dans la section 3.2.2.

Une fois l'extraction des verbes effectuée, nous envisageons dans un deuxième temps d'extraire les cadres de sous-catégorisation des verbes du LST.

### **3.1.2. Extraction des cadres de sous-catégorisation par TAL**

Le travail d'extraction des cadres effectué par Hatier est basé sur le corpus Scientext-SHS annoté aux niveaux morphologique et syntaxique à l'aide de

---

<sup>2</sup> Nous avons effectué le travail du traitement sémantique des verbes du LST dans le cadre du projet TermITH dont l'objectif est d'établir une ressource du LST pour analyser les procédés rhétoriques propres au genre de l'écrit scientifique et pour aider à la rédaction et à la compréhension scientifique. Les traitements sémantiques des autres catégories grammaticales du LST (noms, adjectifs et adverbes) sont effectués en parallèle par d'autres collègues du projet.

l'analyseur en dépendances XIP (Yan & Hatier, 2016 ; Hatier & Yan, 2017). Nous résumons ci-dessous ce travail d'extraction en deux étapes :

Dans un premier temps, pour chaque occurrence verbale, l'ensemble des relations syntaxiques impliquant le verbe étudié est extrait automatiquement à l'aide du corpus analysé en dépendances avec XIP (voir la figure 2-7). Dans un deuxième temps, certaines relations de dépendance sont regroupées sous une même étiquette<sup>3</sup> (ex : passif, complétive), ce qui permet de faire émerger les constructions syntaxiques les plus fréquentes. Il s'agit d'un regroupement des occurrences en convertissant les représentations selon un formalisme proche du modèle CPA.

Considérons la phrase suivante :

Tous les exemples cités **montrent** des orientations explicites. (ling.xml-s6104)

L'analyse syntaxique permet d'obtenir le cadre de sous-catégorisation suivant :

(SUBJ) montrer\_VERB (OBJ)

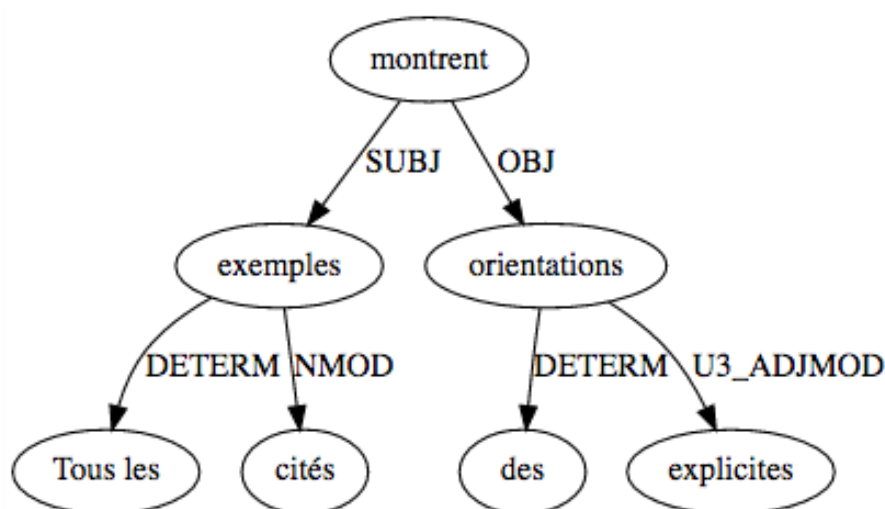


Figure 3-1 : Exemple d'analyse en dépendances

Par ailleurs, pour chaque cadre, les mots en co-occurrence statistiquement significative avec le verbe concerné sont identifiés et calculés. Ces informations nous permettent de dégager, chaque fois que possible, des classes sémantiques préférentiellement associées au verbe pour le cadre en question. Nous avons

<sup>3</sup> Les compléments adverbiaux ne sont pas pris en compte.

également intégré le trait [+hum] lorsque le co-occurent est un des pronoms suivants : *je, on, nous* (le trait [-hum] indique que le co-occurent ne relève pas de ces pronoms). Le tableau 3.1 présente, pour l'exemple du verbe *montrer*, les trois cadres les plus fréquents associés à ses compléments les plus fréquents (le chiffre précédant le co-occurent étant le nombre de fois où le mot apparaît dans cette relation, dans le cadre indiqué).

Cadre de sous-catégorisation - Exemple	Fréquence	Sujet	Objet / Verbe de la complétive
(SUBJ-hum) montrer_VERB(COMPLETIVE_QUE) <i>Nos analyses montrent qu'il y a élaboration des gestes [...]</i>	762	91 : résultat (N) 55 : analyse (N) 41 : étude (N) 27 : travail (N) 17 : tableau (N)	177 : être (V) 53 : pouvoir (V) 45 : avoir (V) 12 : exister (V) 10 : permettre (V)
(SUBJ-hum)montrer VERB(OBJ-hum) <i>Les résultats montrent un effet significatif de la présence de l'appareil [...]</i>	275	29 : analyse (N) 16 : résultat (N) 9 : exemple (N) 8 : figure (N)	13 : importance (N) 13 : existence (N) 10 : différence (N) 10 : effet (N)
(SUBJ+hum) montrer_VERB(COMPLETIVE_QUE) <i>Nous avons montré qu'une combinaison d'indicateurs verbaux [...]</i>	150	103 : nous (P) 39 : on (P) 8 : je (P)	39 : être (V) 16 : pouvoir (V) 8 : permettre (V)

**Tableau 3.1 : Trois cadres de sous-catégorisation du verbe *montrer*<sup>4</sup>**

Il nous est alors possible d'identifier le cadre le plus fréquent (ici 762 occurrences) pour le verbe *montrer*, comprenant un sujet non humain et une complétive introduite par *que*. Sont également renseignés, pour chaque cadre, des exemples ainsi que les co-occurents les plus fréquents. Par exemple, pour le premier cadre, les co-occurents les plus fréquents à la place du sujet sont les noms *résultat*, *analyse*, *étude*, *travail* et *tableau*.

Les résultats de cette phase d'extraction des cadres de sous-catégorisation intègrent ainsi des informations d'ordres syntaxique et lexical. Ces informations sont alors utilisées dans le travail du traitement sémantique des verbes du LST et de la modélisation des patrons verbaux présentés dans la section suivante.

Néanmoins, l'extraction des cadres par les techniques de TAL présente aussi des limites. D'une part, nous avons constaté des erreurs venant de l'analyse syntaxique

<sup>4</sup> Les cadres de sous-catégorisation extraits pour les verbes du LST sont mis en ligne par Hatier et consultables à l'adresse suivante : [http://sylvain.hatier.me/lst/lst\\_csc.php](http://sylvain.hatier.me/lst/lst_csc.php)

automatique. Certaines relations de dépendance sont parfois fausses ou absentes. Par exemple, pour la phrase *cela ne signifie pas seulement que le fait...*, on obtient le cadre (*SUBJ -hum*) *signifier*\_VERBE dans lequel la relation objet est absente. D'autre part, le travail de regroupement de relations de dépendance est à améliorer. Plusieurs occurrences verbales similaires ne sont pas regroupées, ce qui mène à une grande dispersion des cadres et induit par conséquent une diminution de fiabilité de calcul de fréquence.

Compte tenu de ces limites, nous pensons qu'il est indispensable de faire une vérification manuelle à partir des cadres de sous-catégorisation identifiés. Pour cela, nous utilisons le Lexicoscope pour enrichir les informations sur les co-occurrences au niveau des actants tous cadres confondus, et observons, à l'aide des outils Lexicoscope et ScienQuest, les concordances correspondant aux cadres afin de vérifier en corpus leur validité. Par ailleurs, concernant le calcul de fréquence des cadres, étant donné que nous ne pouvons pas avoir une statistique approximative, nous avons choisi de ne pas intégrer le nombre d'occurrences dans notre analyse linguistique de patrons verbaux. L'absence de cette information n'est pas gênante puisque le travail d'extraction des cadres de sous-catégorisation a identifié les constructions syntaxiques les plus fréquentes. Notre objectif est d'analyser les constructions verbales fréquentes et non pas toutes les constructions qui apparaissent dans le corpus, car sur le plan didactique ce sont des éléments intéressants à aborder.

### 3.2. Traitement sémantique des verbes du LST

La liste des verbes du LST obtenue à la suite de l'extraction automatique est une liste de lemmes. Ces verbes peuvent être polysémiques, par exemple le verbe *supposer* renvoie à la formulation d'une hypothèse (*supposons que tout le marché soit effectivement servi...*) ou au fait de produire une conséquence logique (*nos résultats laissent supposer que...*). Il nous semble donc important d'identifier les sens présents pour chaque verbe du LST. Pour ce faire, nous nous basons sur la ressource *Les verbes du français* (LVF). En effet, comme évoqué dans le chapitre I, ce dictionnaire se caractérise d'une part par l'exhaustivité de l'inventaire des emplois des verbes français, et d'autre part par la mise en corrélation entre les constructions syntaxiques

et l'interprétation sémantique dans l'analyse des verbes.

Nous procédons aux traitements sémantiques des verbes du LST qui comportent deux étapes :

1. À partir des cadres de sous-catégorisation, nous repérons les acceptions du LST qui apparaissent dans le corpus Scientext tout en faisant un lien avec la ressource LVF. D'une part, la plupart des acceptions identifiées dans le corpus sont répertoriées dans le LVF ; d'autre part, les acceptions spécifiques à l'écrit scientifique et absentes de la ressource LVF sont créées.
2. En fonction des acceptions repérées, nous proposons un classement sémantique des verbes du LST.

### 3.2.1. Repérage des acceptions avec la ressource LVF

Nous commençons par le repérage des acceptions des verbes du LST présents dans le corpus en nous référant aux entrées verbales répertoriées dans le LVF. Nous travaillons à partir d'un fichier Excel qui conserve tous les verbes extraits (334 au total) et répertoriés dans le LVF. Dans le tableau 3-2, toutes les acceptions considérées comme relevant de l'écrit scientifique sont marquées dans la colonne « LST Oui/Non ».

A MOT	B LST Oui/Non	C Cl. sémantique	D sous-classe	E DOMAINE en clair	F OPÉRATEUR	G CLASSE	H SENS	I PHRASE
aborder 01	N			marine	(#) [grp] lieu qp	S3g.3	accoster	On a~ à la côte, la côte à Marseille.
aborder 02	N			militaire	(#) [ict] qc/qn par bateau	F3a.1	accoster, attaquer	Les pirates a~ le bateau pour le piller.
aborder 03	N			véhicule motorisé	(#) [grp] lieu au début	S3g.3	s'engager dans	On a~ la ligne droite. La voiture a~ le v
aborder 04	O	analyse	#examen	sociologie	(#) [grp] abstrait	S4e.1	commencer, entrepre	On a~ la vie, le latin, un problème. La v
aborder 05	N			sociologie	(#) [grp] qn qp	S1b.1	accoster	On a~ un passant dans la rue pour lui j
aboutir 01	N			véhicule motorisé	(voie) [lc.qp] à fin	L3a.1	se terminer, finir	La route a~ à la mer.
aboutir 02	O	temporalité	fin	temps	(qc) [aux] résultatif A+infii	X1a.5	se conclure à	La discussion a~ à ne rien faire, à un re
aboutir 03	O	temporalité	#fin	temps	(#) [aux] résultatif A+infii	X1a.4	arriver à	On a~ à ne rien faire.
aboutir 04	N			beaux-arts	(#) [mun] surf DE	N3c	recouvrir	Les ouvriers a~ le dôme de feuilles d'o
aboutir 05	N			sociologie	(#) [ag] succès social	H2e.1	réussir	On a~ dans nos recherches.
accéder 01	N			locatif, lieu	(#) [ire] qp	E1d.1	parvenir, arriver	On a~ à la ferme par ce chemin. On a~
accéder 02	N			véhicule motorisé	(voie) [lc.qp] à fin	L3a.1	arriver jusqu'à	Le chemin a~ à la ferme.
accéder 03	N			économie	(qc) [ire] à abstrait	E4c.6	atteindre, arriver à	L'entreprise a~ à un niveau de product
accéder 04	N			psychologie	(#) [loq] oui A qc	C1g	consentir à, accepter	On a~ à sa demande, à ses exigences.
accéder 05	N			droit, administration	(#) [ire] A qc	E2c.1	devenir propriétaire	On a~ à la propriété par les crédits.
accéder 06	O	processus_humain	#obtenir	psychologie	(#) [ire] A abstrait	E2c.1	parvenir, s'élever à	On a~ à la sainteté, au bonheur suprêr
accéder A	O	processus_nonhumain	obtenir				réussir à avoir	Les boxeuses " soft ", en revanche, acc
accentuer 01	N			écriture, écrits	(#) [mun] qc DE accent g	N3a.1	mettre accent graphi	On a~ les mots. Les " e " s'a~ quand ils c
accentuer 02	N			phonétique	(#) [mun] qc DE accent p	N3a.1	mettre accent phoniq	On a~ les syllabes. Les syllabes s'a~ à li
accentuer 03	N			physiologie	(qc) [f.mvt] plus som	M3b.2	souligner, accuser	Les rides a~ son âge. Ses rides s'a~ à l
accentuer 04	O	processus_évolutif	#augmentation	quantitatif	(#) [f.mvt] abstrait vers p	M4b.1	augmenter	On a~ la pression avec cette mesure. L

**Tableau 3.2 : Repérage des acceptions des sens relevant de l'écrit scientifique à partir du modèle lexicographique LVF**

Par exemple, pour le verbe *aborder*, six entrées sont répertoriées dans le LVF,

mais nous n'avons repéré qu'une seule acception (aborder 04) présente dans le corpus. Pour les analyses, nous nous inspirons d'abord des domaines du LVF (voir la colonne E). Nous éliminons dans un premier temps les entrées (aborder 01, 02, 03) appartenant aux domaines sans rapport avec le contexte scientifique comme le domaine de la marine, du militaire, et des véhicules motorisés, etc. Ensuite, nous recensons les acceptions présentes dans le corpus et relevant de l'écrit scientifique<sup>5</sup> à l'aide des cadres de sous-catégorisations et des concordances extraites du corpus. Par exemple, le cadre de sous-catégorisation le plus fréquent du verbe *aborder* se présente comme suit :

(SUBJ-hum) aborder\_VERB(OBJ-hum)

L'observation des exemples du corpus nous révèle que le verbe choisit préférentiellement comme objet les noms *question*, *problème*, *sujet*, *thème*, comme l'illustre l'exemple [1] :

[1] Nous souhaitons **aborder la question** des pathologies sociales et des logiques de réification telle qu'elle a été récemment posée par Axel Honneth. (scinfo\_40)

Nous remarquons que ce sens correspond à l'entrée 04. L'opérateur « [grp] abstrait » veut dire 'saisir quelque chose d'abstrait' (voir la colonne F). Rappelons que dans le chapitre I, nous avons expliqué que l'opérateur consiste en une suite de caractères alphabétiques qui « interprète sémantiquement les schèmes syntaxiques » (Dubois & Dubois-Charlier, 1997, p. 7). Son sens est 'commencer, entreprendre' (voir la colonne H) comme dans la phrase *aborder un problème/la vie*. En conséquence, l'entrée 04 est marquée comme un sens relevant de l'écrit scientifique. Nous appliquons le seuil de fréquence absolue à 20 et celui de dispersion à 3 disciplines. Ainsi, pour qu'une acception soit considérée comme effectivement transdisciplinaire dans le corpus d'analyse, elle doit apparaître au moins 20 fois dans au moins 3 des 10 disciplines. Ce faisant, le repérage des acceptions s'opère dans un mouvement de va-et-vient entre l'observation du corpus, les cadres de sous-catégorisations et les

<sup>5</sup> Parmi les sens présents dans le corpus, nous ne retenons que les sens utilisés dans les activités scientifiques, qu'il s'agisse du positionnement, du raisonnement ou des processus d'écriture. Par exemple, le sens ' manifester un sentiment ' du verbe du LST *exprimer* dans l'exemple *exprimer sa colère* appartenant à la langue générale ne relève pas d'un usage spécifique à l'écrit scientifique. Ainsi, ce sens n'a pas été retenu.

entrées verbales (valeurs sémantiques et constructions syntaxiques) du LVF.

Il est intéressant de constater que les sens des verbes du LST dans l'écrit scientifique relèvent souvent de sens spécifiques ou abstraits qui diffèrent du sens premier indiqué dans les dictionnaires classiques, mais il en est souvent proche. Le sens LST est généralement une spécialisation d'un sens apparaissant dans le LVF. Autrement dit, les sens LST peuvent à la fois se rapprocher des sens du LVF et se spécialiser. Nous présentons ci-dessous deux cas représentatifs rencontrés dans la phase du repérage des acceptions scientifiques.

### 3.2.1.1. Exemple des verbes *appliquer* et *postuler*

Dans l'écrit scientifique, les verbes du LST se caractérisent par leurs comportements sémantique et syntaxique spécifiques liés au genre de l'écrit scientifique. Nous pouvons citer par exemple les cas des constructions au passif et des emplois métonymiques.

D'après Riegel, l'abondance des formes passives dans l'écrit scientifique s'explique par plusieurs fonctions remplies par ces constructions. D'abord, le fait que le complément devienne le sujet permet d'« assurer une progression à thème constant » (Riegel et *al.*, 2009, cité par Fifielska 2015, p. 58). Ensuite, sur le plan stylistique, le passif, et notamment le passif réduit, contribue à « effacer le sujet et alléger la structure phrastique » (*ibid.*). Enfin, le fait d'omettre le premier argument du verbe contribue à « l'impersonnalité du texte technoscientifique » (Kocourek, 1991, cité par Fifielska 2015, p. 58). Prenons l'exemple du verbe *appliquer*. Il est préférentiellement employé à la forme pronominale passive et au passif comme le montre l'exemple [2] :

- [2] Pour cela, nous utilisons les modèles de durée **appliqués** aux périodes de chômage indemnisé. (économie\_34)

Nous mettons en lien ce sens avec l'entrée 10 du verbe *appliquer* qui prend le sens 'adapter' (voir le tableau 3.3).



mot	construction	Opérateur	classe	domaine	Sens	phrase
appliquer 10	T13a0, P30a0	(#) [dat] abstrait A qn	D3b.1	littérature	adapter	On a~ cette remarque à P, à la situation actuelle.

**Tableau 3.3 : L'entrée 10 du verbe *appliquer* dans le LVF**

Le LVF indique deux constructions verbales liées à ce sens : « T13a0 » qui se lit « transitif direct à sujet humain et à objet chose, suivi par un complément introduit par la proposition *à* » et « P30a0 » qui veut dire « construction pronominale à sujet chose et suivi par un complément introduit par la proposition *à* ». Même si le passif n'est pas indiqué dans le LVF, il s'agit bien du même sens. Dans notre tableau d'analyse (tableau 3.2), nous ajoutons la colonne « passif » pour indiquer les cas où le sens s'associe préférentiellement à cette construction dans l'écrit scientifique. Cette information intéressante du point de vue didactique sera ensuite ajoutée dans la modélisation des patrons verbaux. En ce qui concerne les constructions syntaxiques, hormis le passif, nous rencontrons aussi des constructions à l'impératif (ex. : *considérons ces cas*) ou impersonnelles (ex. : *il est montré que*) dont la plupart sont spécifiques ou largement présentes dans l'écrit scientifique. Nous essayons de voir si le sens lié à ces constructions se trouve dans le LVF. Si le sens est présent, nous notons la variation syntaxique, sinon, il nous faut décrire et expliquer le nouveau sens.

Un autre cas intéressant se rapporte aux emplois métonymiques au niveau du sujet. Il s'agit d'un emploi comparable à un emploi avec sujet humain : *l'article propose une étude sur* (= l'auteur propose dans l'article une étude sur). Comme le souligne Tutin (2010a), ces sujets métonymiques courants dans l'écrit scientifique permettent d'éviter la mention explicite de l'agent.

Considérons le verbe *postuler* qui a souvent comme sujet un élément abstrait comme *théorie, approche, recherche, représentation, hypothèse, etc.* (voir l'exemple [3]).

- [3] Notre approche ne **postulant** pas l'existence de structures syntaxiques autonomes par rapport aux termes constitutifs d'un énoncé, le débat ne peut pas être posé dans les mêmes termes.

Il s'agit ici d'un emploi métonymique. En effet, dans l'exemple, ce n'est pas l'approche qui postule, mais c'est la personne qui postule au moyen d'une approche.

De ce fait, nous considérons que le sens de l'exemple [3] peut correspondre à l'entrée 02 du verbe *postuler* dans le LVF, comme dans le tableau 3.4.

mot	construction	Opérateur	classe	domaine	sens	Phrase
postuler 02	T1418	(#) [dic] hypothèse/QUE	C2d.1	psychologie	supposer auprès de	On p~ la bonne foi des joueurs, que l'accord est possible.

**Tableau 3.4 : L'entrée 02 du verbe *postuler* dans le LVF**

L'opérateur « [dic] hypothèse/QUE » veut dire 'communiquer une hypothèse'. La construction « T1418 » indique que le verbe ayant un sujet humain choisit soit un objet inanimé soit une complétive, et peut être suivi par un complément introduit par la préposition *auprès* et un circonstant de manière. Il faut noter que les emplois métonymiques sont largement présents dans l'écrit scientifique et nous ajoutons donc dans notre tableau d'analyse une colonne « métonymie » pour marquer cet emploi spécifique.

### 3.2.1.2. Exemple des verbes *caractériser* et *interroger*

Quand le sens identifié dans le corpus est absent du LVF, nous créons une nouvelle acception pour le verbe dans notre tableau d'analyse. Ceci est souvent le cas quand le verbe accepte un sujet inanimé et qu'il ne s'agit pas d'un emploi métonymique. Nous allons illustrer ce cas avec deux exemples concrets : les verbes *caractériser* et *interroger*.

Pour le verbe *caractériser*, nous relevons deux sens dans le corpus : le premier sens est 'mettre en évidence des traits' où le verbe a un sujet humain, comme dans l'exemple [4] :

- [4] Si on prend en compte le fait que les entités directement touchables sont les substances matérielles, alors on doit **caractériser** les conditions de restriction imposées par ce rôle. (linguistique\_33)

Alors que le deuxième sens est 'constituer le trait distinctif de quelque chose' où le verbe choisit un sujet inanimé comme dans l'exemple [5] :

- [5] Lorsqu'on la situe dans le prolongement des discours du XVII<sup>e</sup> siècle sur le roman,

l'affirmation de la lecture sensible qui **caractérise** la « révolution » de la lecture romanesque au XVIII<sup>e</sup> siècle semble donc moins relever d'une rupture avec un mode de lecture « classique » du roman... (histoire\_50)

En revanche, le LVF ne présente qu'un seul sens pour le verbe *caractériser* (voir le tableau 3.5) :

mot	construction	opérateur	classe	domaine	sens	Phrase
caractériser 01	T1908	ind qc/qn par trait	C3e	titres	typiser	On c~ ce personnage par ce trait. Cette maladie se c~ ainsi.

**Tableau 3.5 : L'entrée 02 du verbe *caractériser* dans le LVF**

L'opérateur « [ind] qc/qn par trait » veut dire « montrer quelque chose ou quelqu'un par un trait » et la construction « T1908 » se lit « transitif direct à sujet humain et à objet chose ou humain, peut être suivi par un circonstant de manière ». Cette entrée verbale rejoint le premier sens. Il semble donc nécessaire de créer une nouvelle acception pour le verbe dans le sens de 'constituer le trait distinctif de quelque chose'.

Considérons maintenant le verbe *interroger* dans les exemples [6] et [7] :

- [6] D'abord, la situation évoquée **interroge** des processus de socialisation complexifiés par ces deux instances autres que la famille que sont le groupe de pairs et les nouveaux supports culturels intégrés à la vie de ces groupes. (scienceseducation.xml-s11233)
- [7] Plus généralement, l'ensemble de ces observations **interroge** la nature des opérations cognitives mises en jeu par la tâche de Running Span. (psycho.xml-s13518)

Dans ces deux exemples, le verbe *interroger* ayant un sujet inanimé prend le sens 'mettre en cause, contester la vérité'. Dans le LVF, les entrées verbales ont un sujet humain et les sens présents impliquent soit un questionnement (voir les entrées 01, 04, et 05) soit une consultation (voir les entrées 02 et 03), étant assez éloignés du sens spécifique trouvé dans le corpus, comme illustré dans le tableau 3.6.

mot	construction	opérateur	classe	domaine	sens	phrase
-----	--------------	-----------	--------	---------	------	--------

interroger 01	T11g0		mand qn SUR	C2i.1	littérature	questionner	On i~ P sur son plan.
interroger 02	T1300 P3000	-	(#) [scrut] qc pour prévoir	P3b.1	physiologie	scruter, consulter	On i~ le ciel pour voir le temps.
interroger 03	T1300 P3000	-	(#) [scrut] ouvrage	P3b.1	littérature	consulter, compulser	On i~ le dictionnaire pour l'orthographe d'un mot.
interroger 04	T11g0		(#) [mand] qn SUR par qc	C2i.1	langue, parole	solliciter du regard	On i~ P du regard sur ce qu'il faut faire.
interroger (s') 05	P1000		ger.mens doute	P1e.1	psychologie	se demander	On s'i~ dans de telles circonstances sur la conduite à avoir.

**Tableau 3.6 : Les cinq entrées du verbe *interroger* dans le LVF**

Ainsi, nous avons créé une nouvelle acception pour le verbe dans le sens de ‘mettre en cause, contester la vérité’ considérant que celui-ci est spécifique à l’écrit scientifique.

Nous avons présenté dans ce qui précède le repérage manuel des acceptions des verbes du LST à travers notamment deux cas représentatifs. À l’issue du repérage, nous obtenons 515 acceptions pour 334 verbes.

La deuxième étape consiste désormais à examiner les traitements sémantiques qui permettront de proposer un classement des acceptions repérées.

### 3.2.2. Classification sémantique

La classification repose sur une typologie à deux niveaux : classes et sous-classes. Celles-ci sont constituées sur la base de propriétés lexico-syntaxiques, à la manière de Dubois & Dubois-Charlier (1997), suite à l’observation et à l’analyse des profils combinatoires et des constructions syntaxiques des verbes du LST dans le corpus Scientext-SHS. La constitution de cette classification nous intéresse sur différents plans : au plan linguistique, elle nous servira de base pour modéliser les patrons lexico-syntaxiques des verbes du LST à la manière du CPA puisque les patrons CPA sont des patrons sémantiques ; au plan didactique, étant donné que notre objectif est la mise en place d’une ressource de patrons par entrée onomasiologique, la classification nous permet de proposer un accès par classe de verbes et fonction rhétorique.

Dans un premier temps, nous nous inspirons des « opérateurs » du LVF. Les éléments primaires des « opérateurs » permettent d'interpréter sémantiquement un verbe en utilisant des fonctions sémantiques telles que « dic » (communiquer quelque chose), « fac » (réaliser une action), « ind » (montrer), etc. À partir de ces fonctions, nous avons commencé à regrouper les acceptions des verbes qui partagent le même code dans la même classe.

Ensuite, nous souhaitons faire émerger des classes lexico-syntaxiques grâce aux constructions syntaxiques. Par exemple, l'opérateur « [ger.mens] QUE » regroupe les acceptions ayant une valeur sémantique 'avoir une telle pensée' et comportant la complétive *que*. Nous obtenons ainsi les verbes suivants qui expriment une opinion neutre : *concevoir, considérer, estimer, juger, supposer*.

Par ailleurs, nous marquons les sous-classes où les acceptions ont un sujet humain avec le signe dièse (ex. : la sous-classe « #comparaison ») pour les distinguer de celles qui ont un sujet non humain lorsqu'il ne s'agit pas d'emplois métonymiques.

Afin d'assurer la cohérence sémantique de la classification, nous consultons également des ressources dictionnaires comme le TLFi<sup>6</sup> pour observer les relations sémantiques entre les verbes du LST telles que la synonymie et l'antonymie. De ce fait, nous considérons que les éléments appartenant à une même classe (Hatier et al., 2016) :

- sont en relation de co-hyponymie
- valident la définition donnée pour la classe
  - ✧ ex : la sous-classe « #comparaison » a pour définition : action scientifique qui vise à rapprocher pour mettre en évidence des rapports de ressemblance ou de différence
- partagent des relations lexico-syntaxiques significatives
  - ✧ ex : les éléments de la sous-classe « #comparaison » (*apparenter, assimiler, comparer, confronter, opposer, rapprocher*) apparaissent préférentiellement dans la construction transitive à sujet humain et à objet inanimé, suivis par un complément prépositionnel introduit par *à* :

<sup>6</sup> Le Trésor de la langue française informatisé (TLFi), consultable en ligne : <http://atilf.atilf.fr/tlf.htm>

*comparer ces résultats à ceux des modèles traditionnels, confronter l'hypothèse à la représentation.*

La polysémie du verbe lui permet d'appartenir à plusieurs classes en fonction de ses propriétés sémantique et syntaxique manifestées. Dans le tableau 3.7., le verbe *expliquer* relève à la fois de la classe « analyse » et de la classe « relation ».

Identifiant	Catégorie	Classe	Sous-classe	Définition	Exemple
expliquer-1	Verbe	analyse	Interprétation	faire comprendre	L'auteur explique en détail la réflexion qui l'a amenée à sélectionner Twitter. (scinfo.xml-s10822)
expliquer-2	Verbe	relation	implication_cause	constituer une des causes de	Plusieurs facteurs expliquent cette situation. (scienceseducation.xml-s11822)

**Tableau 3.7 : Acceptions du verbe *expliquer* dans le corpus Scientext-SHS**

Le classement comporte 12 grandes classes étiquetées comme suit : ancrage\_construction, observation, analyse, communication, état, processus évolutif, temporalité, processus humain, processus non humain, relation, processus cognitif\_psycho et modalité. Ces classes contiennent 102 sous-classes. Les classes et les sous-classes des verbes du LST sont détaillées dans l'Annexe 3.

Il faut noter que les classifications sémantiques manuelles sur les quatre catégories (nom, adverbe, adjectif et verbe) ont été effectuées concomitamment et de façon collaborative<sup>7</sup> pour garantir l'homogénéité transcatégorielle dans le cadre du projet Termith. Ainsi, certaines classes couvrent plusieurs catégories étudiées, par exemple la sous-classe « description » couvre le verbe *présenter-2* dans le sens de 'montrer, exposer' (*les résultats sont présentés dans le tableau*) et le nom *présentation-1* dans le sens de 'fait de montrer' (*cette présentation des différents types de théories pragmatiques*).

Le travail de classification des verbes du LST a permis de mettre en évidence le lien étroit entre les constructions verbales et leur interprétation sémantique. Chaque sous-classe regroupe des verbes ayant des propriétés syntaxiques et sémantiques

<sup>7</sup> Un travail d'harmonisation a été effectué par Magdalena Augustyn entre les classifications des catégories grammaticales (adjectifs, noms, adverbes et verbes). Nous la remercions également pour son aide dans la construction des classes des verbes.

similaires. Quelques exemples de sous-classes sont présentés dans le tableau suivant :

Classe	Sous-classe	Définition	Acception verbale	Constructions syntaxiques
<b>ancrage_constructi on</b> (construction du modèle d'analyse : problématique, hypothèses, conceptualisation, définition, réflexion, références)	#fondement	Action scientifique qui vise à prendre qqch pour base, pour principe fondamental et de démontrer une chose en justifiant la démonstration par ces principes (des faits, des théories, des preuves, une documentation) qui soutiennent l'hypothèse.	ancrer, appuyer_1, appuyer_s'_2, baser_1, baser_se_2, établir_1, fonder, inspirer_s'_2, reposer_se_2	N[sujet]+V+N[objet]+ sur/de/dans N[objet]+(se)V + sur/de N[objet]+(passif)V+ sur/de/dans
	#finalité	Avoir en vue et mettre en œuvre son action dans un but déterminé	attacher_s'_2, consacrer_se_1, efforcer_se, employer_s'_1, envisager_2, proposer_se_2, tendre_1, tenter, viser_1	N[sujet]+(se)V + à/de + V/N N[sujet] +V+ N
<b>relation</b> Classe regroupant les verbes renvoyant aux différents types de relation entre entités ou exprimant une mise en relation effectuée selon différents critères.	#association	Action de mettre en relation, réunir en un tout des éléments divers qui se complètent et produisent un effet unique.	accompagner_1, articuler, associer, combiner_1, conjuguer, corréler, lier, rattacher, relier	N[sujet]+V+N[objet1]+à/avec +N[objet2]
	Source	Relation mettant en lien l'origine, la raison de qqch et le résultat	découler, dépendre, hériter, procéder_2, provenir, ressortir_1, résulter	N[sujet]+V+de N

**Tableau 3.8 : Exemples de sous-classes**

Une fois la classification des verbes effectuée, nous obtenons une ressource des verbes du LST dont les entrées lexicales sont des acceptions, avec pour informations, la classe, la sous-classe et une courte définition. Il nous reste à enrichir cette ressource avec l'analyse des patrons lexico-syntaxiques par classe de verbes.

Nous allons présenter dans la section suivante la modélisation des patrons fondée sur le modèle CPA en précisant les points qui différencient notre travail de celui de Hanks. Rappelons que notre objectif est plus spécifique, car lié à l'écrit scientifique et qu'il s'inscrit dans une perspective didactique d'aide à la rédaction scientifique.

### 3.3. Utilisation du modèle CPA dans l'écrit scientifique

Rappelons que notre choix du modèle CPA se justifie par les caractéristiques du modèle. D'abord, le CPA est présenté comme une méthode destinée à décrire le comportement lexico-syntaxique du verbe à travers l'analyse des patrons. Les patrons CPA sont considérés comme des patrons de sens. Comme le souligne EL Maarouf (2011, p. 89), l'avantage de ce modèle est qu'il permet de « décrire le sens d'un verbe en contexte en fonction des patrons principaux dans lesquels il est employé ». Dans un deuxième temps, il utilise des catégories sémantiques nommées « types sémantiques » pour caractériser les arguments nominaux qui apparaissent dans un patron (*cf.* Chapitre I), ce qui affine l'analyse du sens. Ainsi, en identifiant les patrons principaux d'un verbe, nous prenons en compte non seulement les constructions syntaxiques récurrentes, mais aussi les collocations/co-occurrences qui entourent le verbe. Enfin, le CPA accorde un rôle primordial aux corpus et aux contextes. Ce faisant, les patrons représentent les emplois usuels et réels du mot dont les apprenants ont besoin. En vue de faciliter les tâches rédactionnelles liées aux écrits universitaires, il nous semble donc primordial de dégager chaque sens de mot en contexte, ce qui permettra ensuite la désambiguïsation lexicale.

Étant donné que notre objectif est de relever les emplois usuels du verbe dans le domaine de l'écrit scientifique avec une visée didactique, il nous a semblé essentiel d'adapter le modèle CPA à nos besoins. En fait, la modélisation des patrons des verbes du LST doit déboucher sur une ressource de patrons (sous format lexicographique) à laquelle l'enseignant peut avoir accès pour l'aide à l'écrit scientifique. Dans la section suivante, nous présenterons notre modélisation des patrons des verbes du LST en réfléchissant sur la façon de présenter cette ressource de patrons. Nous aborderons les questions essentielles suivantes :

1. Comment typer les arguments nominaux dans un patron donné ?
2. Comment présenter les sens du verbe ?
3. Comment présenter les constructions syntaxiques spécifiques à l'écrit scientifique ?



4. Comment représenter les patrons des verbes du LST avec un objectif didactique ?

### 3.3.1. Attribution des types/rôles sémantiques

#### 3.3.1.1. Description sémantique au niveau des arguments d'un patron

Une caractéristique importante du modèle CPA est d'identifier les préférences sémantiques au niveau des arguments d'un patron donné afin de désambiguïser les sens du verbe (*cf.* Chapitre I). Pour ce faire, il faut repérer les unités lexicales apparaissant dans une position syntaxique donnée et leur attribuer un trait sémantique. Les unités lexicales récurrentes peuvent être considérées comme un ensemble de co-occurents qu'on peut regrouper sous des étiquettes sémantiques homogènes. El Maarouf (2011) a expliqué que Hanks utilise le SketchEngine qui génère les co-occurents d'un verbe donné automatiquement. En analysant les co-occurents d'un verbe dans une position donnée, Hanks choisit le(s) type(s) sémantique(s) approprié(s) dans le patron. Par exemple, pour le verbe *achieve*, le SketchEngine génère les principaux co-occurents en position objet comme suit: *objective* (353 occurrences), *goal* (269), *result* (247) *aim* (171), *sucesss* (170). Au vue de ces co-occurents, le type sémantique `[[State_of_Affairs]]` a été établi pour représenter la préférence sémantique en position objet, comme montré dans le patron suivant :

Pattern : `[[Human | Institution]] achieve [[State_of_Affairs = Desired Objective]]` (taux de fréquence=100%)

Implicature : `[[Human | Institution]] successfully causes [[State_of_Affairs = Desired Objective]] to happen`

Dans le CPA, les types sémantiques sont établis au fur et à mesure de la construction des patrons. Par ailleurs, en fonction de l'interprétation sémantique de ces collocatifs en contexte, le rôle sémantique `[[Desired Objective]]` a été combiné au type sémantique `[[State_of_Affairs]]` afin d'affiner la description sémantique.

Comme dans le CPA, nous ajoutons des types sémantiques dans la construction d'un patron. Dans un premier temps, nous nous appuyons sur les co-occurents

extraits automatiquement lors de l'étape de l'extraction des cadres de sous-catégorisations (voir le tableau 3.1 présenté précédemment). Nous avons également consulté l'outil Lexicoscope pour repérer les co-occurents d'un verbe dans une position syntaxique donnée. Par exemple, pour le cadre suivant du verbe *rejeter*, nous avons relevé que les principaux co-occurents en position objet direct sont *hypothèse, idée, modèle, solution, théorie*, etc. Ici, le verbe signifie « refuser ».

(SUBJ +hum)rejeter\_VERB(OBJ -hum)

Les exemples suivants illustrent cette préférence sémantique.

- [8] Loin de **rejeter les théories classiques** de l'anthropologie culturelle du XXe siècle, **on** montre que l'obsession de l'authenticité culturelle n'est pas nouvelle et que l'éloge du métissage n'en est que le revers. (sciencespolitiques.xml-s23551)
- [9] **Certains rejettent** catégoriquement **l'idée d'une préposition vide** ou, plus généralement, d'un mot dépourvu de sens. (ling.xml-s7582)
- [10] **L'hypothèse de marchés complets** étant **rejetée**, **nous** nous intéressons aux alternatives expliquant comment les ménages parviennent à lisser intertemporellement leur consommation dans un environnement de marchés incomplets. (eco.xml-s16133)

Afin de relever les co-occurents significatifs, nous fixons le seuil de fréquence à cinq occurrences. Dans un deuxième temps, au vu des co-occurents relevés, nous attribuons un trait sémantique en nous référant aux classes/sous-classes des noms du LST établies dans le cadre du projet TermITH (Hatier, 2016). La classification sémantique des noms du LST nous permet de typer les arguments nominaux dans un patron donné. Ce choix peut être expliqué par le fait que les noms du LST ont été extraits de la même manière que les verbes du LST et qu'une désambiguïsation sémantique a été effectuée. Nous présentons dans le tableau 3.9 les principales classes/sous-classes qui sont utilisées comme types sémantiques dans notre analyse de verbes d'analyse et de positionnement (*cf.* Chapitre IV).

classe	sous-classe	Exemple
support (communication)	document, graphique, section	<i>article, tableau, figure, annexe, introduction 1</i>
espace	domaine, localisation, orientation	<i>contexte, niveau, direction, voie 1</i>
qualité (état)	caractéristique, complexité,	<i>propriété, qualité 1, difficulté,</i>

	importance, manière	<i>obstacle, problème_1, fonction, importance, façon, manière</i>
objet scientifique	cas, donnée, finalité, <b>méthode</b> , thème, <b>thèse</b>	<i>cas, phénomène, corpus, donnée, but, objectif, approche, démarche, problème_2, <b>idée</b>, <b>hypothèse</b></i>
humain	collectivité, individu	<i>communauté, auteur, chercheur</i>
processus cognitif	choix, constat, démonstration, description, évaluation, examen	<i>adoption, identification, observation, démonstration, illustration_2, analyse, étude</i>
processus évolutif	augmentation, baisse, changement	<i>augmentation, développement</i>
processus humain	action, usage	<i>utilisation, activité, travail</i>
quantité	mesure	<i>différence_1, pourcentage, résultat_1</i>
relation	association, implication, opposition,	<i>lien, relation, conséquence, effet, divergence, résultat_2</i>

**Tableau 3.9 : Les types sémantiques utilisés dans l'analyse des patrons**

Cette liste des types sémantiques pourrait être complétée et enrichie par d'autres types sémantiques au fur et à mesure de la construction des patrons. Il faut noter que la plupart des unités lexicales relevées peuvent être catégorisées avec les classes/sous-classes des noms du LST. Cependant, lorsque la majorité des occurrences n'appartient pas aux noms du LST, nous choisissons de créer un type sémantique pour les représenter.

Reprenons le cadre du verbe *rejeter* présenté plus haut, les principaux co-occurents en position objet direct (*hypothèse, idée, modèle, solution, théorie*) renvoient à un objet abstrait qui construit l'activité scientifique. Nous pouvons les regrouper sous les types sémantiques [[**méthode|thèse**]] (surlignés en jaune dans le tableau 3.9). Le patron peut donc être présenté comme suit :

[[**humain**]] rejette [[**méthode|thèse**]]

Contrairement au type sémantique qui représente le caractère intrinsèque du mot, le rôle sémantique est déterminé par le contexte. Considérons les exemples suivants du verbe *expliquer* :

[11] Cependant, nous pouvons supposer que ce n'est pas le seul **facteur** qui explique **le succès** continu des politiciens d'origine étrangère. (sciencespolitiques.xml-s20969)

[12] **Les difficultés financières de la ville (Zawatka - Gerlach, 2005) et des principes d'aménagement en conséquence plus pragmatiques** [15] expliquent en partie **cette différence**. (geo.xml-s1427)

- [13] *La dépense importante que ces spectacles entraînent* explique peut-être *leur absence* en certains diocèses [151]. (histoire.xml-s16856)

Ici, le verbe prend le sens de ‘constituer une des causes de’. Les noms *facteur*, *difficulté* et *dépense* en position sujet peuvent être catégorisés comme [[élément]] explicatif. Les noms *succès*, *différence* et *absence* en position objet peuvent être regroupés sous les types sémantiques [[cas|relation]]. En analysant une variété des occurrences au niveau des arguments du verbe, nous pouvons proposer le patron suivant :

[[élément=cause]] explique [[cas|relation=conséquence]]

Nous avons associé le rôle sémantique [[cause]] au type sémantique [[élément]] et [[conséquence]] aux types [[cas|relation]]. Ces rôles sont donc liés au sens du verbe *expliquer*. L’attribution des rôles sémantiques permet de mettre en évidence le sens du verbe et contribue donc à la désambiguïsation lexicale, notamment lorsque deux patrons ont le même type sémantique. Considérons les deux verbes *confronter* et *améliorer* qui acceptent le même type sémantique [[situation]] dans les patrons suivants :

[[humain]] se confronte à [[situation=négatif]] (co-occurents : *situation*, *difficulté*, *problème*, *échec*, *environnement*, etc.)

[[humain]] améliore [[situation]] (co-occurents : *condition*, *confort*, *situation*, *environnement*, *bien-être*)

Ici, le type sémantique [[situation]] peut être défini comme l’ensemble des conditions matérielles ou morales dans lesquelles se trouve une personne. Dans ces deux patrons, les unités lexicales en position objet sont du même type sémantique et pourtant les co-occurents sont différents. En effet, d’une part, notre corpus est limité. D’autre part, cette différence est liée au sens de ces deux verbes. Le verbe *confronter* implique une signification négative ‘se trouve contraint de faire face à quelque chose’ et il mobilise plus souvent des mots à polarité négative (*difficulté*, *échec*, etc.). Le verbe *améliorer* veut dire ‘changer en mieux’ et est souvent entouré des mots tels que *condition*, *confort*, *situation*, *environnement*, etc. Nous avons donc associé le rôle sémantique [[négatif]] au type sémantique [[situation]] dans le patron du verbe

*confronter*. D'ailleurs, il semble important de souligner qu'en fonction du sens des verbes ou du contexte, les unités lexicales qui entrent dans un type sémantique ne sont pas transférables à tous les verbes qui peuvent accepter le même type.

Enfin, en combinant les types et les rôles sémantiques, nous parvenons à décrire les unités lexicales au niveau des arguments d'un patron donné, faisant ressortir le sens du verbe.

### 3.3.1.2. Difficultés rencontrées

Lors de l'attribution des types sémantiques, nous avons été confrontée à la difficulté de classer certaines unités lexicales parmi les types sémantiques définis dans le projet TermITH. En effet, certains noms combinés au verbe échappent à une classification satisfaisante en raison des limites du type sémantique qui nous semble convenir au mieux ou des préférences lexico-syntaxiques du verbe.

Par exemple, le verbe *mobiliser* est identifié comme un verbe de l'activité humaine et le seul sens repéré est 'mettre en œuvre des moyens'. Ce verbe accepte comme sujet un nom humain et prend comme objet quelque chose d'abstrait. Les principaux co-occurents en position objet sont *ressource*, *savoir*, *stratégie*, *approche*, *procédé*, *instrument*, *notion* et *outil*. La majorité des occurrences en position objet renvoient donc soit à une méthode ou à un instrument mis en œuvre pour l'activité scientifique, soit au savoir relatif au domaine de la connaissance. Il est donc possible de les représenter avec les types sémantiques [[méthode|savoir|instrument]], Nous présentons le patron comme suit :

[[humain]] mobilise [[méthode|savoir|instrument]]

Ainsi, l'exemple [14] peut correspondre à ce patron :

[14] En mobilisant **une approche intégrative** mêlant la problématique du lien social sur réseaux à celle de la construction identitaire, **nous** tenterons de faire une place à l'activité des sujets. (scinfo.xml-s9155)

Comme on peut le constater, le co-occurent *ressource* du verbe *mobiliser* n'est pas considéré comme un nom du LST et ne peut entrer dans aucun des types

sémantiques proposés ci-dessus. Ce mot échappe donc à notre classification sémantique.

Par ailleurs, si l'on considère le type sémantique [[instrument]], on trouve dans la liste des noms du LST les noms *moyen, instrument, outil, technique, matériel, enquête, entretien, questionnaire*, etc., mais il nous semble que certains noms comme *enquête* et *entretien* sont rarement combinés avec le verbe *mobiliser*.

Face aux difficultés liées à l'attribution des types sémantiques, une des solutions serait de lister les principaux co-occurrents dans une position syntaxique donnée et de les associer au type sémantique. L'intégration des co-occurrents permet de mettre en lumière les relations collocationnelles entre le verbe et les mots qui l'entourent. En outre, rappelons que l'identification des préférences sémantiques au niveau des arguments d'un patron donné consiste à mettre en évidence le sens du patron. Les patrons des verbes du LST sont des patrons prototypiques et il ne faut pas qu'ils deviennent des modèles à suivre impérativement. En fin de compte, il est impossible que les types/rôles sémantiques décrivent toutes les possibilités du mot puisque d'une part, notre corpus est limité et que d'autre part, la langue est un système évolutif et de ce fait, l'interaction entre les sens du verbe et ses arguments est imprévisible :

The interaction between the meanings of V and that of N may cause the semantic type of N to be “stretched” or “shrunk” in composition, so that for a given verb type selectional requirement, no precise prediction seems possible as to what words make up the corresponding lexical set in context. (Jezek & Frontini, 2010, p. 3)

Ainsi, pendant l'attribution des types sémantiques, le choix d'un type sémantique est souvent le résultat d'un compromis. D'un côté, nous tenons à proposer des catégories sémantiques qui incluent la majorité des unités lexicales ; de l'autre, nous ne listons pas toutes les possibilités, ce qui risquerait d'alourdir la lecture du patron.

### 3.3.2. Vers des patrons syntaxiques spécifiques au genre

Dans le modèle CPA, un patron représente une structure syntaxique, souvent la forme standard du verbe, en principe la voix active, et correspond à un sens contextualisé du mot. En revanche, il est important de souligner qu'un patron du

verbe peut « correspondre à de nombreuses réalisations (variations) en contexte, mais conserve le même sens » (El Maarouf, 2011, p.71). Autrement dit, un sens du verbe peut être réalisé à travers plusieurs patrons (variations syntaxiques).

Selon Hanks, le sens du patron prime sur les variations syntaxiques et de ce fait, les variations syntaxiques qui n'ont pas d'influence sur le sens du verbe (par exemple, la phrase à la forme passive) ne sont pas traitées comme un patron à part. On se démarque de Hanks sur ce point, puisque nous identifions les patrons verbaux dans le genre de l'écrit scientifique. Il apparaît important de prendre en compte les variations syntaxiques dans la modélisation des patrons dans la mesure où les verbes du LST se caractérisent par des constructions spécifiques, importantes à souligner au plan didactique. À partir des cadres de sous-catégorisations extraits, nous pouvons constituer des patrons syntaxiques. Nous présentons ci-dessous quelques cas de variations syntaxiques rencontrées dans notre analyse avec notamment l'exemple du verbe *expliquer*. Dans l'écrit scientifique, ce verbe possède une large gamme de structures syntaxiques.

### 3.3.2.1. Les constructions passives

Dans la partie 3.2.1.2., nous avons évoqué l'abondance des constructions passives dans l'écrit scientifique, qui s'explique par plusieurs fonctions remplies par ces constructions, en particulier le fait de mettre en position de sujet les résultats scientifiques, produisant par conséquent un effet objectivant. Dans la modélisation des patrons, nous ne traitons pas la variation passive comme un patron à part. Ici, la variation au passif inclut le passif réduit (*résultat obtenu*) ou le passif périphrastique en *être* + participe passé (*être déterminé par*). Nous n'intégrons la variation au passif que lorsqu'un verbe s'emploie exclusivement dans la forme passive, ce qui peut caractériser un sens particulier du patron. Si les formes passives sont fréquentes, nous choisissons de marquer leurs présences sous forme de commentaires (*cf.* Chapitre IV). Prenons l'exemple du verbe *expliquer*, dont un des cadres extraits dans la forme passive est :

(OBJ -hum)(PASSIF)expliquer\_VERB

Ce cadre peut s'actualiser dans les deux exemples ci-dessous :

[15] Remarquons que certains **phénomènes** peuvent déjà être **expliqués** à la lueur de ce que nous avons dit précédemment. (ling.xml-s985)

[16] Si les élèves des deux groupes se soumettent à des tests de connaissance en fin de chaque année, il sera possible de réaliser les deux **comparaisons expliquées** ci-dessus...(scienceseducation.xml-s8089)

Ici, la forme passive ne contribue pas à la caractérisation d'un sens particulier du patron et nous avons donc choisi de l'associer à la phrase active (voir l'exemple 17) :

[17] **On** peut ainsi **expliquer l'attrait contemporain** pour la notion de réseau en santé, dans un contexte de difficulté de financement des dépenses de santé, par la nécessité d'améliorer la performance du système de santé, à moyens inchangés, voire même réduits. (economie\_21)

Dans les exemples ci-dessus, le verbe *expliquer* signifie 'faire comprendre'. Ainsi, le patron pourrait être présenté comme suit :

[[**humain=auteur, chercheur**]] (pouvoir) explique [[**cas|processus|relation**]]

### 3.3.2.2. Les différents types de constructions impersonnelles

Comme les constructions passives, plusieurs types de constructions impersonnelles sont relevés dans notre corpus : contextes impersonnels (*il est intéressant de constater*), constructions avec verbes impersonnels introducteurs (*il faut noter que*), constructions impersonnelles (*il a été montré que*). Ces constructions se caractérisent par l'existence d'un sujet implicitement associé à l'agent du procès, ce qui permet de mettre en avant les activités et les faits scientifiques plutôt que les agents qui effectuent ces activités. Ainsi, ces constructions remplissent-elles souvent des fonctions similaires aux constructions passives. D'une part, elles effacent l'agent du procès et contribuent ainsi à « exprimer une information subjective d'une manière qui semble objective » (Fifielska, 2015, p. 51), et d'autre part, elles provoquent un changement de la structure thématique de l'énoncé et servent donc à introduire un nouveau sujet. Il est intéressant de constater que ces différents types de constructions impersonnelles sont souvent associées à différentes fonctions rhétoriques, par exemple, exprimer la nécessité de l'étude (*il est important de*), la possibilité du fait (*il*



*semble que*), introduire une interprétation (*il en découle que*), etc. Compte tenu de la diversité des fonctions rhétoriques que jouent ces constructions, il s'avère important de les organiser selon leurs fonctions rhétoriques manifestées. Ceci pourrait constituer une entrée onomasiologique intéressante pour la rédaction scientifique. En revanche, la présence de ces constructions ne donne pas lieu à la création d'un nouveau patron. Nous choisissons de les marquer sous forme de commentaires pour ne pas alourdir la représentation des patrons (cf. Chapitre IV). Par exemple, nous avons relevé des contextes impersonnels (*il est à noter que, il convient de noter que, il est intéressant de noter que*) et des constructions avec verbes impersonnels introducteurs pour le verbe *noter* dans notre corpus (*il faut noter que*). Nous les associons au patron principal ci-dessous qui s'actualise dans l'exemple 18 :

[humain=auteur, chercheur] (pouvoir) note (adverbe discursif) [complétive que]

[18] On peut noter que le facteur d'impact de certaines revues en libre accès s'envole.  
(scinfo.xml-s13373)

Le verbe *noter* prend le sens 'introduire une information supplémentaire et invite le lecteur à y prêter l'attention'.

### 3.3.2.3. Les constructions en *comme*

Les constructions en *comme* sont très productives dans l'écrit scientifique. Ces marqueurs introduisent « des comparatives méta-énonciatives » (Debaisieux et Martin 2010, p. 321 ; cité par Grossmann 2014a, p. 764). Comme le souligne Grossmann (*ibid.*), ces constructions ont pour fonctions principales de faire un rappel ou un renvoi afin de guider le lecteur (*comme évoqué précédemment, comme illustre le tableau suivant*), et d'introduire une référence afin de renforcer l'argumentation (*comme le fait remarquer X*). Il est intéressant de remarquer que ces constructions sont plus contraintes que les constructions passives, étant strictement associées à certains groupes de verbes tels que les verbes de constat (*observer*), verbes de démonstration (*montrer*), verbes de proposition (*suggérer*), verbes de communication (*évoquer*) et verbes de révélation (*illustrer*), etc. Les rôles rhétoriques ainsi que le nombre limité de ces constructions nous amènent donc à les faire figurer dans la représentation des patrons. Sur le plan didactique, la maîtrise de ces constructions

s'avère importante. Considérons ci-dessous les trois exemples de constructions en *comme* du verbe *expliquer* :

- [19] Comme l'explique **Nadelmann**, l'OCRTIS [17] était dans les années 1970 - 1980 le correspondant privilégié de la DEA et filtrait les relations de ce dernier avec les autres forces de police [18]. (sciencespolitiques.xml-s2078)
- [20] Vient aussi s'y greffer le phénomène des accès aux articles par liens extérieurs, comme **nous** l'avons expliqué plus haut. (scinfo.xml-s13120)
- [21] L'ambiguïté du message est aussi illustrée par une forme en quelque sorte intermédiaire entre la ruine et le mémorial, assez répandue, comme l'explique **Jacques Sauvageot** : « Parmi ces commémorations, la pratique la plus courante consiste en... » (Sauvageot, 1995).

Ici, le verbe *expliquer* signifie 'faire comprendre'. Ces exemples pourraient entrer dans le patron suivant, celui-ci se situe donc à côté des autres patrons ayant le même sens :

comme l'explique [[**humain=auteur, chercheur**]]

#### 3.3.2.4. Les constructions périphrastiques causatives '*permettre de*', '*amener à*' et '*conduire à*'

Un autre type de constructions récurrentes relevé pendant notre analyse des verbes se rapporte aux constructions périphrastiques causatives (*permettre de/amener à/conduire à*+verbe à l'infinitif). Ces constructions comportent un sujet inanimé qui renvoie souvent aux activités ou aux objets scientifiques causateurs et un complément d'objet indirect introduit par la préposition *de* ou *à*, celui-ci renvoyant à un agent humain (sujet du verbe à l'infinitif) est souvent implicite. Ces constructions impliquent « une relation de cause à effet » (Bak et Novakova, 2013, p. 101), comme en témoignent les exemples 22, 23 et 24 :

- [22] Ces multiples **méthodes** ont **permis** au final **d'identifier** 110 blogs et 42 sites natifs de l'internet répondant aux critères précédemment évoqués. (scinfo.xml-s11284)
- [23] **Les observations** conduites aux Cosmonautes **amènent à s'interroger**, en

conclusion, sur le caractère potentiellement heuristique d'approches...  
(sciencespolitiques.xml-s20709)

- [24] **La comparaison** de ces résultats **conduit** immédiatement à **énoncer** la proposition suivante. (eco.xml-s11441)

Le sujet inanimé, souvent un objet (*méthode*) ou une activité scientifique (*observation, comparaison*) (cause A) détermine les actions à réaliser (effet B) dans la recherche de l'auteur (ou des auteurs). Il est intéressant de constater que les verbes à l'infinitif relèvent souvent des verbes qui décrivent le processus d'analyse ou d'interprétation (*identifier, analyser, proposer, choisir, considérer*, etc.). La fonction principale de ces constructions est donc d'apporter des explications pour justifier le bien-fondé de la recherche scientifique, impliquant ainsi un raisonnement causal. Par ailleurs, nous pensons que ces constructions constituent également un moyen d'évitement de la mention de l'auteur, permettant au lecteur de se centrer sur les événements et la relation causale, ce qui provoque un effet de l'objectivité des recherches scientifiques. Il faut noter que la construction *permettre de+verbe* à l'infinitif est beaucoup plus fréquente que les deux autres dans notre corpus. Ainsi, lors de la modélisation des patrons, nous choisissons de l'intégrer dans la représentation des patrons alors que les deux autres seront soulignées sous forme de commentaires. Par exemple, l'exemple 25 pourrait se réaliser dans le patron suivant :

[[**objet scientifique |activité scientifique=cause**]] permet d'expliquer  
[[cas|processus|relation=conséquence]]

- [25] **L'analyse** des origines des instruments d'action publique environnementale de l'UE (voir tableau 2) **permet**, dans un premier temps, **d'expliquer** de manière détaillée **l'existence** d'un système décisionnel à plusieurs niveaux, ainsi que **la capacité limitée** de l'UE à l'innovation instrumentale. (sciencespolitiques.xml-s17221)

Il est important de souligner que les cas de variations syntaxiques présentés ci-dessus sont étroitement liés au choix des verbes à analyser dans le chapitre IV. En élargissant l'étude des verbes du LST, il y aura peut-être d'autres constructions syntaxiques spécifiques à prendre en compte dans l'analyse des patrons.

### 3.3.3. Représentation des patrons par entrée onomasiologique

Dans le chapitre I, nous avons expliqué que le modèle CPA a été appliqué dans la construction du dictionnaire *Pattern Dictionary of English Verbs (PDEV)* pour identifier les patrons d'emploi usuel des verbes anglais de la langue générale. Ce dictionnaire suit une représentation classique avec comme entrée le verbe. Chaque verbe regroupe différents patrons qui sont associés aux sens du verbe. Par exemple, le verbe *explain* présente cinq patrons différents, ce qui témoigne de la polysémie de ce verbe (voir la figure 3-2).

**Figure 3-2 : Les patrons du verbe *explain* dans le PDEV**

On peut remarquer que les patrons et les sens du verbe sont présentés de manière linéaire en fonction de leur fréquence. Dans notre étude, l'analyse des patrons du verbe vise à proposer à l'enseignant du FLE une ressource des patrons exploitable pour la classe. De ce fait, l'entrée par le verbe et les patrons en premier ne nous

#### PDEV: explain

[Access full data](#)

 Displayed here are **All patterns** .  
 Other options: [Phrasal verbs](#) |

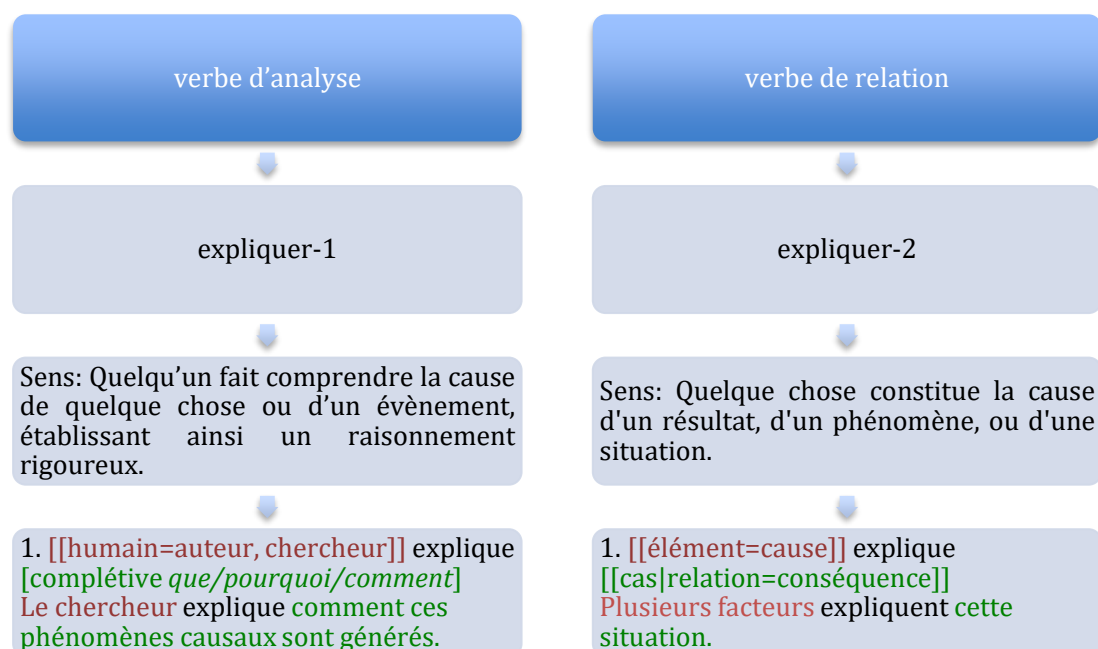
 samplesize: **500**  
 patterns: **5**

1	Pattern: <b>Human</b> or <b>Institution</b> or <b>Document</b> or <b>Proposition</b> or <b>Eventuality 1</b> <b>explains</b> <b>Eventuality 2</b> (in terms of <b>Eventuality 3</b> ) (to <b>Human 2</b> ) Implicature: <b>Human</b> or <b>Institution</b> formulates (in conversation with <b>Human 2</b> or in a <b>Document</b> intended for <b>Human 2</b> to read) a <b>Proposition</b> concerning the cause or effect of an <b>Eventuality 2</b> + Example: It seemed impossible to <b>explain</b> the similarities by similar learning experiences.	58.6%	<a href="#">...More data</a> FrameNet
2	Pattern: <b>Human</b> or <b>Institution</b> or <b>Document</b> or <b>Proposition</b> <b>explains</b> (to <b>Human 2</b> ) Implicature: <b>Human</b> or <b>Institution</b> formulates <b>Proposition</b> in <b>Document</b> concerning the cause or effect of an <b>Eventuality</b> that is expressed in WH- [CLAUSE] Example: To <b>explain</b> why students need critical thinking skills.	18.2%	<a href="#">...More data</a> FrameNet
3	Pattern: <b>Human 1</b> or <b>Document</b> <b>explains</b> <b>THAT-CLAUSE</b> (to <b>Human 2</b> ) Implicature: <b>Human 1</b> or <b>Document</b> tells <b>Human 2</b> <b>THAT</b> [CLAUSE], as a way of accounting for <b>Eventuality</b> (often implied rather than stated explicitly) Example: He <b>explained</b> that union membership was an outdated rule.	11.2%	<a href="#">...More data</a> FrameNet
4	Pattern: <b>Human</b> <b>explains</b> <b>QUOTE</b> Implicature: <b>Human</b> says [QUOTE] as a way of accounting for <b>Eventuality</b> (often implied rather than stated explicitly) Example: She <b>explains</b> : 'Most adults' respiratory pattern is upset.	9.8%	<a href="#">...More data</a> FrameNet
5	Pattern: PHRASAL VERB. <b>Human 1</b> <b>explains</b> <b>Eventuality</b> <b>away</b> Implicature: <b>Human 1</b> formulates <b>Proposition</b> as a way of accounting for <b>Eventuality</b> = <b>Inconvenient</b> Example: This staggering figure could not easily be <b>explained</b> away.	0.8%	<a href="#">...More data</a>

semble pas pertinente puisque l'enseignant de langue ne fonctionne pas de façon linéaire dans l'enseignement. Ainsi, nous choisissons de proposer deux types de représentation : l'un se base sur la classification sémantique des verbes du LST, l'autre sur les fonctions rhétoriques des constructions verbales. Autrement dit, l'accès aux patrons peut se faire soit par classe du verbe soit par fonction rhétorique.

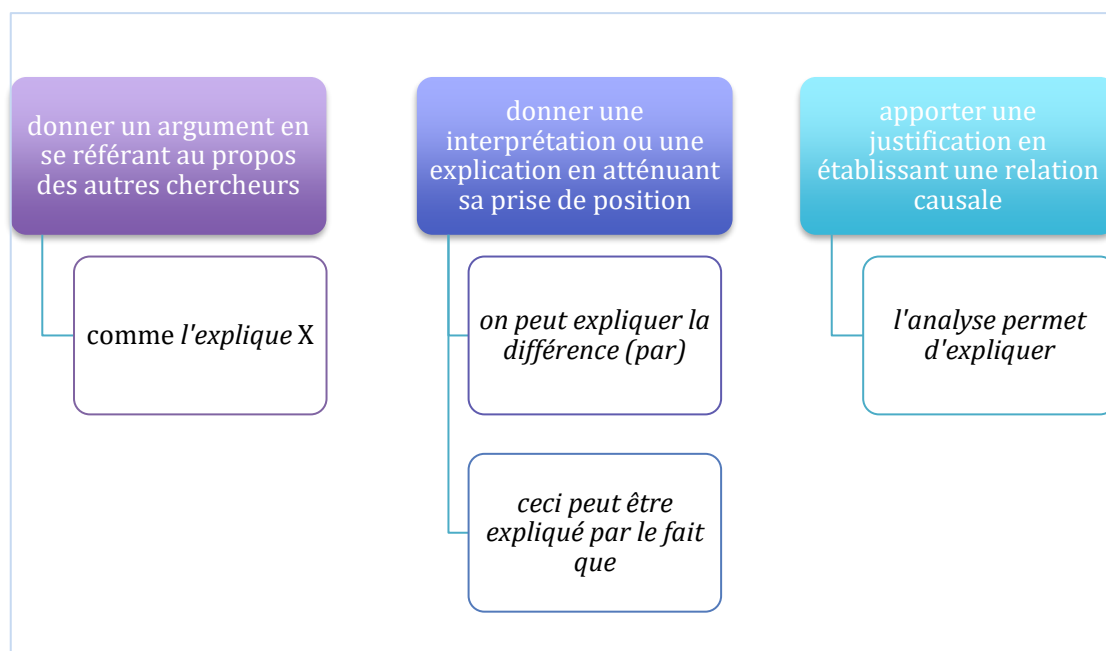
D'abord, l'entrée par classe du verbe permet d'accéder directement à l'acception verbale en question et d'étudier les verbes synonymes qui partagent des propriétés sémantiques et syntaxiques similaires. Rappelons que dans le tableau 3.7., nous avons

relevé deux sens du verbe *expliquer*, relevant à la fois de la classe « analyse » et de la classe « relation ». Les patrons identifiés du verbe *expliquer* seront donc présentés respectivement dans ces deux classes, comme l'illustre la figure 3-3 :



**Figure 3-3 : Représentation des patrons du verbe *expliquer* avec l'entrée par classe**

Deuxièmement, l'entrée par fonction rhétorique permet d'accéder à différentes constructions verbales de la même fonction et faciliterait par conséquent l'encodage et l'enseignement de l'écrit. Dans le chapitre IV, les patrons et les constructions verbales analysés seront organisés selon leur fonction rhétorique. Compte tenu des cas de variation syntaxique mentionnés ci-dessus, nous illustrons dans la figure 3-4 l'entrée par fonction rhétorique avec l'exemple du verbe *expliquer*.



**Figure 3-4 : Représentation des patrons et des constructions verbales avec l'entrée par fonction rhétorique-exemple du verbe *expliquer***

Après avoir présenté la modélisation des patrons verbaux adaptée à notre étude avec un objectif didactique, nous résumons dans la section suivante les points communs et les différences entre le modèle CPA et notre adaptation.

### 3.3.4. Récapitulatif des points communs et des différences

De manière générale, nous nous démarquons des patrons CPA du fait que nous ne cherchons pas à identifier les patrons d'emploi usuels des verbes dans la langue générale, mais nous nous intéressons à un genre de texte particulier qui est l'écrit scientifique. Notre objectif spécifique implique une adaptation du modèle CPA tant au niveau de la méthode d'analyse qu'au niveau de l'analyse des patrons. Nous proposons ci-dessous un tableau récapitulatif qui résume les points communs et différents entre le modèle CPA et notre adaptation.

		CPA	Adaptation aux verbes du LST
Au niveau de la méthode d'analyse	points communs	1. une méthode d'analyse de corpus 2. une méthode destinée à décrire les sens du verbe en contexte	
	points différents	1. « corpus-driven » 2. repérage manuel des patrons	1. « corpus-driven » et « corpus-based » 2. repérage semi-automatique des patrons à l'aide des techniques de TAL
Au niveau de l'analyse des patrons	points communs	1. chaque patron est associé à un sens 2. utilisation des catégories sémantiques pour typer les arguments nominaux	
	points différents	1. types sémantiques établis dans une ontologie sémantique 2. prise en compte des variations syntaxiques seulement si celles-ci provoquent un changement de sens 3. un sens est représenté par un patron	1. utilisation des noms du LST pour typer les arguments nominaux 2. prise en compte des variations syntaxiques spécifiques à l'écrit scientifique 3. un sens peut être représenté par plusieurs patrons

**Tableau 3.10 : Comparaison entre le modèle CPA et notre adaptation**

### 3.4. Bilan

Dans ce chapitre, nous avons expliqué la procédure d'extraction des verbes du LST et des cadres de sous-catégorisation par les techniques de TAL. Ce travail effectué par Hatier nous permet d'effectuer un repérage semi-automatique des patrons lexico-syntaxiques : l'extraction automatique des cadres a réduit le temps de l'analyse des occurrences verbales, mais la vérification manuelle a été nécessaire pour enrichir et valider les données obtenues.

Ensuite, nous avons montré que la plupart des acceptions des verbes du LST relevées dans le corpus Scientext-SHS peuvent correspondre aux entrées verbales dans le LVF. En revanche, des spécificités propres à l'écrit scientifique, liées aux collocations et aux constructions syntaxiques nécessitent d'être prises en compte. Le travail du repérage a abouti à une classification des acceptions verbales qui facilite l'analyse des patrons verbaux. Enfin, en adaptant le modèle CPA à un corpus

spécialisé, nous avons expliqué différents choix d'analyse pour mieux saisir les spécificités sémantiques et syntaxiques des verbes du LST. Les problèmes principaux posés par la modélisation des patrons à partir d'un corpus spécialisé concernent d'une part la façon de typer les arguments nominaux, et d'autre part, le traitement des variations syntaxiques.

Après avoir expliqué la méthode de modélisation des patrons, nous procédons à l'analyse des verbes du LST dans le chapitre suivant. Dans le cadre de cette étude, nous nous intéressons à deux classes de verbes, les verbes d'analyse et les verbes de positionnement.



---

## Chapitre IV.

# Analyse linguistique des verbes du LST

---

4.1. INTERET POUR LES VERBES CHOISIS .....	126
4.1.1. Choix des verbes d'analyse.....	126
4.1.2. Choix des verbes de positionnement .....	126
4.2. PARAMETRES DE L'ANALYSE .....	127
4.3. LES VERBES D'ANALYSE .....	128
4.3.1. Verbes de comparaison.....	129
4.3.2. Verbes de constat.....	136
4.3.3. Verbes de description.....	143
4.3.4. Verbes d'examen.....	148
4.3.5. Verbes d'interprétation.....	154
4.3.6. Récapitulatif.....	160
4.4. LES VERBES DE POSITIONNEMENT .....	163
4.4.1. Verbes d'opinion.....	165
4.4.2. Verbes de démonstration.....	170
4.4.3. Verbes de mise en relief.....	176
4.4.4. Verbes de choix.....	182
4.4.5. Verbes d'hypothèse .....	188
4.4.6. Verbes de proposition .....	194
4.4.7. Verbes d'inférence .....	201
4.4.8. Récapitulatif.....	206
4.5. BILAN.....	209

---

Pour faire des constructions verbales du LST un objet d'enseignement et d'apprentissage en tant que tel, nous avons adopté la méthodologie de l'analyse des patrons, inspirée du modèle CPA et proposé une classification des verbes du LST basée sur leurs propriétés lexico-syntaxiques. À partir de ce travail, ce chapitre est consacré aux analyses syntaxiques et sémantiques des patrons verbaux qui apporteront des éclairages aux éléments à prendre en compte dans l'E/A des patrons. Parmi les verbes du LST retenus dans ce travail, nous avons choisi de nous concentrer sur deux classes de verbes : les verbes d'analyse et les verbes de positionnement. Ces verbes ont été choisis parce que d'une part, ils jouent des rôles importants dans l'écrit scientifique, et d'autre part, ils semblent poser problème aux étudiants dans la rédaction académique.

## 4.1. Intérêt pour les verbes choisis

### 4.1.1. Choix des verbes d'analyse

Notre choix des verbes d'analyse est d'abord lié à leurs rôles importants dans l'énonciation. L'étude des verbes permet de mettre en évidence les rôles de chercheur assumés par le scripteur (Fløttum et al., 2006). Dans l'activité d'analyse, l'auteur/chercheur présente les démarches scientifiques et le processus de construction des connaissances dans la discipline.

Du point de vue didactique, l'enseignement de l'écrit scientifique consiste d'abord à amener les apprenants à « manier les notions/concepts de manière cohérente et rigoureuse » (Deschepper & Thyron, 2008, p. 67) à travers notamment l'établissement du constat scientifique, l'analyse des données ainsi que l'interprétation des résultats. Il s'agit pour eux d'« entrer dans une démarche scientifique et s'affirmer comme sujets apprentis-chercheurs dans un discours qui doit mettre à distance leur propre subjectivité. » (Lafont-Terranova & Niwese, 2012, p. 118-119). En ce sens, la maîtrise des verbes d'analyse permet de développer la compétence d'associer des connaissances, des objets et des concepts (Parpette & Mangiante, 2010).

### 4.1.2. Choix des verbes de positionnement

Les verbes de positionnement jouent un rôle important dans l'argumentation, caractéristique clé des écrits scientifiques (Hyland, 2005b ; Rinck, 2006) centrée sur « la relation qui relie le lecteur et le scripteur, et celle qui s'établit entre le scripteur et les auteurs qu'il cite » (Henderson, 2014, p. 145). Dans l'écrit scientifique, le scripteur marque son attitude et cherche à convaincre afin d'être accepté en tant que membre d'une communauté de discours (Swales, 1990), ce qui conduit à la mise en fonctionnement du positionnement (le rôle d'argumentateur).

En effet, le positionnement est un champ sémantique déjà abordé dans le cadre du projet Scientext et d'autres études (Rinck, 2006) (Tutin et al., 2009) puisque « tout

chercheur doit apprendre à se démarquer/se positionner par rapport à sa communauté de pairs » (Cavalla, 2015, p. 97).

Sur le plan didactique, ces verbes font partie des verbes sous-employés par les apprenants (Granger & Paquot, 2009b). Cavalla (2015) a souligné que les apprenants ont des difficultés à se positionner dans leur écrit académique, souvent à cause de lacunes lexicales pour le faire de façon appropriée (Chavez, 2008, Mroué, 2012). Par ailleurs, les apprenants ont du mal à s'affirmer comme auteurs dans leur discours « par le biais du travail langagier sur l'objet concerné et sur les discours existants à son propos » et la prise de position se caractérise souvent par l'affirmation de leur subjectivité. (Deschepper & Thyron, 2008, p. 74). Ainsi, nous pensons qu'il est important d'aider les apprenants à développer la compétence d'argumentation. La maîtrise des verbes de positionnement constitue donc un objectif pédagogique crucial.

Dans les sections 4.3 et 4.4, notre approche lexicale et phraséologique analyse les constructions verbales et les fonctions des verbes d'analyse et de positionnement, et par conséquent amène à comprendre l'importance de la construction d'une identité de chercheur/argumentateur dans l'écrit scientifique. Avant de procéder à l'analyse, il nous semble nécessaire d'explicitier les paramètres de notre analyse.

## **4.2. Paramètres de l'analyse**

Nous commençons par définir chaque sous-classe de verbes en insistant sur leurs propriétés communes. Nous nous référons également au dictionnaire TLFi pour proposer une définition générale pour chaque sous-classe de verbes. Dans un deuxième temps, nos analyses sémantiques et syntaxiques des patrons verbaux vont s'articuler autour de trois paramètres :

1. les constructions syntaxiques et valeurs sémantiques
2. les pronoms sujets mettant en jeu le scripteur
3. les arguments nominaux

En premier lieu, nous choisissons d'aborder les constructions syntaxiques en

expliquant les valeurs sémantiques associées. Ce choix se justifie par le fait que chaque construction syntaxique possède une valeur spécifique et a tendance à apparaître dans un contexte discursif précis. Par exemple, la construction en *comme* revêt souvent une valeur évidentielle et métatextuelle (*cf.* Chapitre III), utilisée concomitamment avec l'argumentation ou la déclaration d'un fait. Par ailleurs, le traitement combiné des aspects syntaxiques et sémantiques permet de faire ressortir les différentes nuances de sens et de mieux comprendre le fonctionnement spécifique de chaque verbe au sein du discours scientifique.

Ensuite, si l'on souhaite étudier la dimension énonciative, il est nécessaire de prendre en compte les manifestations pronominales de l'auteur. Concernant les verbes d'analyse et de positionnement, l'auteur est l'agent des activités d'analyse et de prise de position. L'étude des pronoms sujets mettant en jeu l'auteur nous permet d'interroger sur les rôles de l'auteur dans ces activités scientifiques et les relations qu'ils établissent vis-à-vis du lecteur et des auteurs qu'il cite.

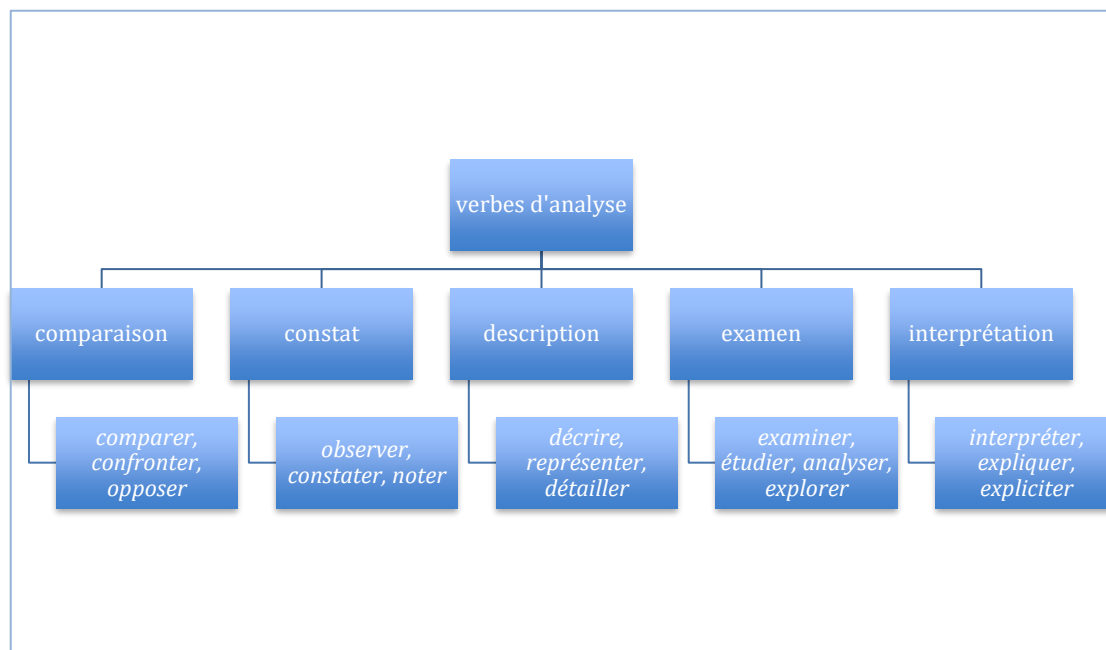
Enfin, au niveau des arguments nominaux, nous axons nos analyses sur les sujets métonymiques et les objets nominaux. D'une part, les emplois métonymiques du type *cette étude* consistent en « des stratégies d'évitement de la mention explicite de l'agent » (Tutin, 2010a, p, 20). Nous pensons que les sujets métonymiques, fréquents dans l'écrit scientifique, sont révélateurs de la voix de l'auteur et qu'ils doivent être pris en compte dans notre analyse des verbes. D'autre part, l'analyse des objets nous permet de relever les collocations du type V+N.

À l'issue des analyses de chaque sous-classe, nous présentons les patrons sous forme de tableau. Celui-ci constitue donc la ressource finale destinée aux enseignants et fera l'objet de l'exploitation à titre d'exemple dans le chapitre V.

### **4.3. Les verbes d'analyse**

Comme nous l'avons déjà souligné, les verbes d'analyse, utilisés dans notre corpus d'écrit scientifique, remplissent un rôle d'analyse dans les activités liées au processus de recherche : description des données, analyse des relations entre les variables, comparaison des résultats attendus, interprétation des écarts et des résultats.

Ces verbes relèvent de la grande classe conceptuelle « la scientificité » (Pecman, 2007) qui regroupe les concepts renvoyant à « une action, un objet, une entité, une abstraction, une qualité, etc., de nature purement scientifique » (*ibid.*, 293). À notre connaissance, ces verbes n'ont pas encore fait l'objet d'une étude systématique dans l'écrit scientifique. La figure ci-dessous présente les cinq sous-classes que nous allons étudier.



**Figure 4-1 : La classification des verbes d'analyse à étudier**

Sous chaque classe, nous avons choisi d'étudier les verbes fréquents et représentatifs. Par ailleurs, l'observation du corpus nous permet de voir que ces verbes présentent des constructions syntaxiques variées et qu'ils sont associés à des fonctions rhétoriques importantes. Nous allons procéder à l'analyse des verbes par sous-classe.

### 4.3.1. Verbes de comparaison

Les verbes de comparaison relevés sont interprétés sémantiquement dans le LVF comme des verbes qui relient deux éléments pour les différencier ou les identifier. Ils sont utilisés pour 'mettre en évidence des rapports de ressemblance ou de différence' (d'après le TLFi) entre deux objets d'étude. Ces derniers sont « à certains égards identiques, semblables ou dissemblables » (Siepmann, 2007, p. 125). Dans l'activité

scientifique, la comparaison constitue une méthode d'analyse importante. On compare ses propres méthodes ou ses résultats à ceux des pairs en vue de dégager les traits ou les spécificités de l'étude. Trois verbes prototypiques sont sélectionnés : *comparer* (524), *confronter* (383), *opposer* (704). Notons que seul le verbe *comparer* est monosémique, non ambigu et est toujours employé dans cette acception.

### a. Constructions syntaxiques et valeurs sémantiques

Dans cette sous-classe, les trois verbes prennent comme sujet un nom humain et objet un nom inanimé. Deux cas apparaissent : 1) l'objet inanimé consiste en deux ou plusieurs éléments à comparer [1] 2) l'objet inanimé comporte le premier élément à comparer et celui-ci est suivi par le complément prépositionnel *à* qui introduit le deuxième élément [2].

[1] Ce processus constituerait, ainsi, un réseau qui faciliterait la récupération de la cible. Pour mettre ceci en évidence, les auteurs **comparent des mots** de fréquence d'usage différente. (ling.xml-s10940)

[2] Malgré ces limites, nous avons voulu **confronter nos résultats** sur la ségrégation académique et sociale **aux données** disponibles concernant l'organisation des systèmes éducatifs européens essentiellement recueillies par Eurydice (2005) [11] et présentées dans les Chiffres clés de l'éducation. (scienceseducation.xml-s160)

Dans ces deux exemples ci-dessous, l'introduction de la comparaison est justifiée vis-à-vis de l'objectif (*pour mettre ceci en évidence*) et de l'intention. Ici, les deux éléments mis en parallèle sont censés partager à la fois des points communs et différents. Quant au verbe *opposer*, les éléments à comparer sont présentés comme totalement différents [3] :

[3] **En opposant** ainsi **la liberté du corps** au **conformisme**, l'émancipation à l'aliénation, Pascal Ory livre une vision quelque peu téléologique, dans l'idée de critiquer ceux qui verraient dans le bronzage « de nouvelles ruses du contrôle social » ou une « décadence ». (histoire.xml-s22649)

Le verbe *confronter* englobe à la fois les idées de l'examen et de la comparaison. Il signifie 'examiner un objet scientifique ou un fait dans un esprit de comparaison en

vue de dégager des résultats'. Dans l'exemple [4], on pourrait dire que les données sont examinées par rapport aux catégories proposées.

- [4] Essayons à présent de **confronter nos données aux catégories proposées** par Philippe Descola. (anthropologie.xml-s12772)

Si l'on observe les constructions les plus fréquentes, on s'aperçoit que le gérondif est fréquemment employé pour établir la comparaison : *en comparant* (64/524), *en confrontant* (12/383), *en opposant* (22/704). Ici, le gérondif introduit un moyen, étant souvent accompagné de l'expression des résultats dans le co-texte qui suit [5] :

- [5] **En comparant ces résultats** avec ceux d'autres projets similaires (le compte Flickr de la Bibliothèque du Congrès et sur celui du projet PhotosNormandie), nous avons fait apparaître que cette faible efficacité n'est pas structurelle...(scinfo.xml-s10941)

L'emploi du gérondif permet donc d'alléger la structure phrastique et d'établir une relation de l'effet (résultat) à la cause (activité de la comparaison). Nous avons également relevé un nombre non négligeable de constructions au passif pour le verbe *comparer* (94/524), y compris le passif réduit [6] et le passif comportant le modalisateur *pouvoir* [7] :

- [6] Globalement, tous ces résultats sont conformes à Barsalou et Sewell (1984), à cette nuance près qu'on observe une différence selon le type de catégorie : les catégories d'activité **comparées aux** catégories taxinomiques déterminent un accord et une variation généralement plus important. (psycho\_43)

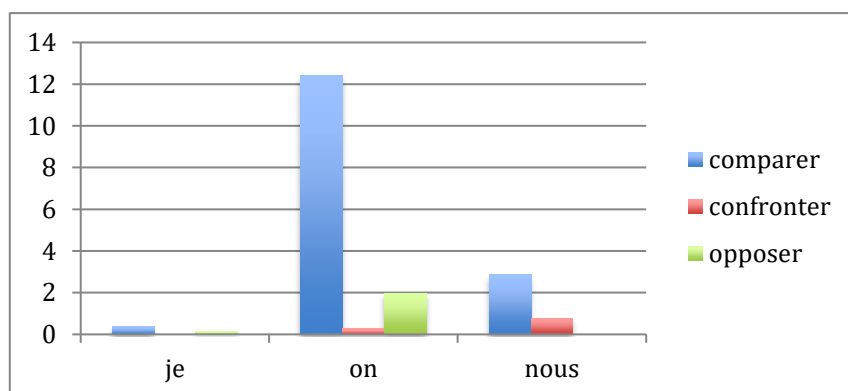
- [7] Ce classement **peut être comparé** aux résultats d'une enquête menée dans quatre villes libres (kreisfreie Städte) de différentes régions d'Allemagne (Bamberg, Bonn, Ludwigshafen et Wiesbaden) auprès des conseillers municipaux [34]. (sciencespolitiques.xml-s1509)

Le fait de mettre en position de sujet les éléments à comparer produit un effet objectivant dans l'analyse comparative. D'une manière générale, le verbe *comparer* est le plus productif en termes de constructions : *il est intéressant de comparer* (13/524), *lorsque/si l'on compare* (30/524), *permettre de comparer* (21/524). D'abord,

le contexte impersonnel *il est intéressant de comparer* permet de diriger l'attention du lecteur vers des phénomènes essentiels. Ensuite, les constructions *lorsque/si l'on compare* semble inviter le lecteur dans l'activité de la comparaison. Enfin, la construction causative *permettre de comparer* contribue à établir la relation de l'effet (l'activité de la comparaison) à la cause.

### b. Pronoms sujets mettant en jeu l'auteur

Le nombre non négligeable des constructions au passif et au gérondif nous amène à penser que la mention de l'auteur dans l'étude de comparaison est souvent implicite. Nous pouvons donc postuler que les verbes de comparaison ne sont pas fréquemment associés aux pronoms sujets mettant en jeu l'auteur. La figure 4-2 illustre l'association entre les trois verbes de comparaison et les pronoms sujets *je*, *on* et *nous* :



**Figure 4-2 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de comparaison)**

Conformément à notre hypothèse, les trois pronoms sujets ne sont pas fréquents. Il est intéressant de constater que le verbe *comparer* s'associe davantage au pronom *on* qui est souvent suivi par le modalisateur *pouvoir* [8] :

- [8] **On peut comparer** l'étude de la discrimination à l'étude du crime, la criminologie, qui pose des problèmes analogues. Or les criminologues et sociologues de la justice ne font pas des enquêtes sur les personnes susceptibles d'être victimes de crime, mais sur les crimes, sur leurs auteurs et leurs victimes. (anthropologie.xml-s4978)

En plus, comme indiqué plus haut, il apparaît souvent dans la subordonnée de



condition *lorsque/si l'on compare*, ce qui justifie un emploi inclusif du pronom *on* (Fløttum *et al.*, 2006) où l'auteur souhaite ouvrir un dialogue avec la communauté du lecteur, l'incluant ainsi dans le déroulement de l'analyse comparative [9] :

- [9] **Quand on compare** les rapports de Flaubert et de George Sand avec leur éditeur commun, on constate que c'est la femme de lettres qui a conquis non seulement l'estime, mais l'admiration la plus débordante de l'éditeur. (histoire.xml-s11174)

Les pronoms *on* et *nous* sont peu fréquents avec les verbes *confronter* et *opposer* dans le sens de comparaison, ce qui peut en partie s'expliquer par la polysémie de ces deux verbes. Par contre, nous avons relevé quelques cas dans lesquels le verbe *opposer* est associé au sujet du pronom *on*. Il est intéressant de constater que la valeur de *on* est différente de celle de l'exemple [9]. Dans l'exemple [10], le pronom *on* représente les autres chercheurs. Cet emploi du pronom *on* figure parmi les six valeurs définies par Fløttum *et al.* (2006)<sup>1</sup>.

- [10] **On oppose** traditionnellement la défense des libertés individuelles et la libre circulation (le droit des individus sur lequel repose l'essentiel de notre édifice juridique) au principe d'un contrôle du centre vers la périphérie. Mais les choses ne sont pas si simples...(sociologie.xml-s10435)

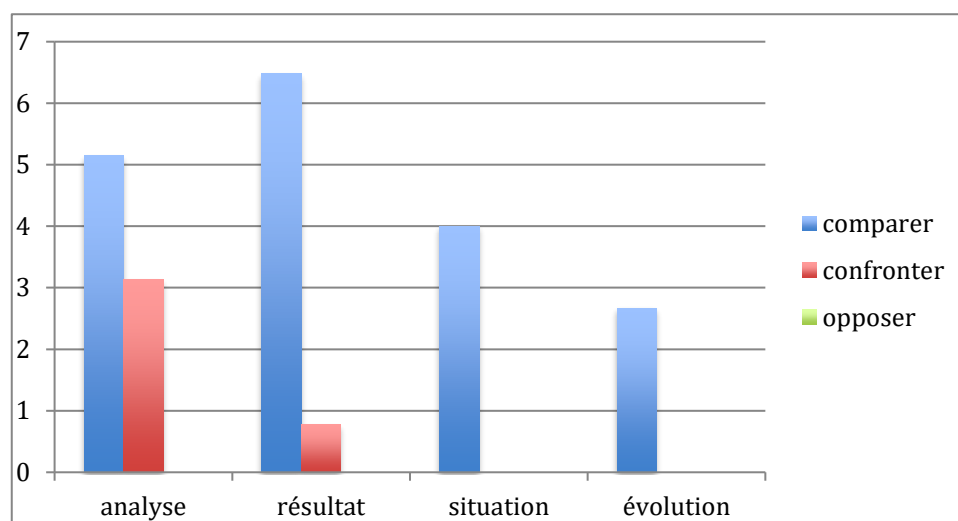
Cet exemple montre que le verbe *opposer* à polarité négative n'est pas fréquemment utilisé par les auteurs pour décrire leurs activités de la comparaison. Dans l'acception de 'comparer', il nous semble que le verbe *comparer* a plus de fonctions que les deux autres verbes.

### c. Arguments nominaux

Au niveau du sujet, ces trois verbes choisissent un sujet humain (souvent l'auteur). Quant à l'objet, les éléments à comparer présentent une grande variété. Il

<sup>1</sup> Selon Fløttum *et al.*, il existe six valeurs de *on* dans l'écrit scientifique, situées soit du côté personnel soit du côté indéfini. Ces six valeurs sont présentées comme suit : « ON1 représente l'auteur correspondant à *je/nous* ; ON2 remplace l'auteur et les lecteurs correspondant à *je/nous+les* lecteurs ; ON3 renvoie à l'auteur et la communauté discursive limitée correspondant à *je/nous+mes/nos* collègues ; ON4 représente l'auteur et la communauté « non limitée » correspondant à *je/nous+* « tout le monde » ; ON5 représente les lecteurs ; ON6 remplace les autres correspondant aux autres chercheurs ».

apparaît que le verbe *comparer* est le plus productif. La figure 4-3 donne un aperçu des noms objets les plus fréquents :



**Figure 4-3 : Les noms objets des verbes de comparaison (en %)**

Pour les verbes *comparer* et *confronter*, les noms en commun relevés sont *analyse* et *résultat*. Le verbe *comparer* peut aussi s'associer aux noms abstraits qui se rapportent à une activité ou à un cas (*situation*) [11] tandis que le verbe *confronter* ne s'associe à ces noms que dans le sens de 's'affronter à'.

[11] On montre, enfin, de quelle manière le prix garanti optimal associé aux contrats proposés pour le gel de terre permet de limiter le coût social des rentes d'information **lorsque l'on compare les situations** d'information parfaite et d'information asymétrique. (eco.xml-s17461)

Concernant le verbe *opposer*, le nombre faible des noms objets ne nous permet pas de proposer un aperçu général. Comme signalé plus haut, il semble moins utilisé par les auteurs dans leur analyse comparative du fait de son sémantisme négatif mettant en contraste deux éléments.

Pour les trois verbes de comparaison étudiés, les objets à comparer sont souvent le thème de la phrase, surtout si l'on considère le nombre récurrent de constructions au passif. Si ces trois verbes partagent des constructions similaires sur le plan syntaxique, ils se distinguent par contre en termes de sémantisme des objets. Contrairement au verbe *opposer* davantage mobilisé pour présenter deux choses totalement différentes, *comparer* et *confronter* semblent être davantage employés dans

le processus de comparaison, à côté du contexte exprimant l'objectif, la cause ou l'effet de la comparaison, bien qu'il s'agisse de simples tendances.

Enfin, les patrons lexico-syntaxiques<sup>2</sup> sont présentés dans le tableau 4.1 :

comparer	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] (pouvoir) compare [[objet scientifique processus cognitif événement=pluriel]] Ex : <i>Enfin, <b>on peut comparer les choix</b> du capital et du prix dans l'incertain dans le cas certain.</i> (eco.xml-s126)</li> <li>2. [[humain=auteur, chercheur]] (pouvoir) compare [[objet scientifique 1 processus cognitif 1 événement 1]] à [[objet scientifique 2 processus cognitif 2 événement 2]] Ex : <i>L'hypothèse de la politisation progressive dans l'opinion de l'indice longitudinal de tolérance est d'autant plus plausible qu'<b>on compare cet indice à un indice de tolérance</b> envers les juifs.</i> (sciencespolitiques.xml-s19566)</li> <li>3. [[méthode processus cognitif]] permet de comparer [[objet scientifique processus cognitif événement=pluriel]] Ex : <i>En effet, la réflexion sur les tarifs donne lieu à des calculs élaborés <b>permettant de comparer</b> différentes options...</i> (histoire.xml-s25084)</li> </ol> <p>Lexset [[objet scientifique processus cognitif événement]] : <i>résultat, analyse, évolution, situation</i> Sens : Quelqu'un met en parallèle deux objets d'étude afin d'en dégager les traits ou les spécificités.</p>
confronter	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] confronte [[objet scientifique processus cognitif=pluriel]] Ex : <i>En <b>confrontant</b> diverses données, Serhij Plokhij put montrer récemment que...</i> (histoire.xml-s18440)</li> <li>2. [[humain=auteur, chercheur]] confronte [[objet scientifique 1 processus cognitif 1]] à [[objet scientifique 2 processus cognitif 2]] Ex : <i><b>Nous ne confronterons pas ici notre analyse à celles de ces auteurs</b> afin de ne pas sortir du cadre de cet article.</i> (ling.xml-s8791)</li> </ol> <p>Lexset [[objet scientifique processus cognitif]] : <i>résultat, analyse, donnée</i> Sens : Quelqu'un met un résultat ou une analyse face à un autre (une autre) afin de mettre en évidence les rapports de ressemblance ou de différence sur lesquels fonder son opinion.</p>
opposer	<ol style="list-style-type: none"> <li>1. [[humain]] oppose [[entité abstraite événement=pluriel]] Ex : <i><b>On peut schématiquement opposer deux lectures</b> de l'histoire de la discipline.</i> (anthropologie.xml-s13635)</li> <li>2. [[humain]] oppose [[entité abstraite 1 événement 1]] à [[entité abstraite 2 événement 2]] Ex : <i><b>Les chercheurs</b> en relations internationales ont longtemps campé sur deux positions antagoniques, <b>opposant la démarche</b> stato-centrée des intergouvernementalistes <b>aux hypothèses</b> néo-fonctionnalistes.</i> (sciencespolitiques.xml-s17603)</li> </ol> <p>Sens : Quelqu'un met en contraste deux éléments qui sont présentés comme totalement différents, de manière à faire valoir les ressemblances ou les</p>

<sup>2</sup> Ici, nous ne pouvons pas proposer le nombre des patrons identifiés puisqu'à cause des limites de l'extraction des cadres de sous-catégorisation (cf. Chapitre III), il est impossible de regrouper les cadres par acception manuellement. Rappelons cependant que nous appliquons le seuil de fréquence absolue à 20 et celui de dispersion à 3 disciplines pour qu'une acception soit considérée comme effectivement présente et transdisciplinaire dans le corpus d'analyse. Ces critères nous semblent donc suffisants pour relever les patrons les plus fréquents.

	différences sur lesquelles se fonde son opinion.
Valeurs sémantiques	valeur comparative
Routines	<i>il est intéressant de comparer</i> (13/524) <i>lorsque/si l'on compare</i> (30/524)
Commentaires	Le passif et le gérondif sont présents, le pronom <i>on</i> est fréquent pour le verbe <i>comparer</i> .

**Tableau 4.1 : Patrons lexico-syntaxiques des verbes de comparaison**

### 4.3.2. Verbes de constat

Dans l'écrit scientifique, la découverte et l'interprétation s'établissent souvent à partir de l'observation des données et des faits. Celle-ci est donc une étape centrale de toute recherche scientifique. Les trois verbes relevés sont *observer* (1428), *constater* (619) et *noter* (250). Dans le LVF, ces verbes sont classés comme des verbes psychologiques et partagent le sens de base ' manifester telle activité réflexive sur qn/qc'. Dans l'activité scientifique, ces verbes ont pour principale fonction de 's'assurer, au moyen d'observations scientifiques, de la réalité d'un fait' (d'après le TLFi). En d'autres termes, il s'agit bien « de mettre en scène une opération intellectuelle d'analyse de données plus que de simple perception ou d'activité graphique. » (Grossmann, 2014a, p. 760)

#### a. Constructions syntaxiques et valeurs sémantiques

Pour les trois verbes de constat, les structures privilégiées sont la construction avec complétive conjonctive (*on constate que*) et celle avec un complément nominal (*on observe un effet*). Ces deux constructions, apparemment préférées dans le constat des activités scientifiques, acceptent un sujet humain et prennent le sens de « 'prendre connaissance de la réalité d'un fait au moyen d'observation scientifique' » (Grossmann, 2014b, p. 86). D'une manière générale, les verbes *constater* et *noter* s'associent fréquemment à la construction en complétive alors que le verbe *observer* a une préférence pour le complément nominal.

Dans la construction avec la complétive conjonctive, les verbes de constat permettent d'introduire une information ou un résultat obtenu après l'analyse des

données [12], ou à l'aide des éléments péri-textuels (*tableau, figure, etc.*) [13] :

[12] Des huit interventions que j'analyse, **j'ai constaté que** ces constructions peuvent se diviser en deux catégories selon le rôle qu'elles jouent au sein de ces interventions. (ling.xml-s11253)

[13] Cependant, on peut voir dans la partie basse du tableau que cette tendance n'est pas confirmée dans deux cas : l'arabe littéraire et le français. **On peut observer que** non seulement les chiffres de la colonne des "non" en production sont les mêmes que ceux des "non" en réception, mais qu'en outre...(linguistique\_522)

Concernant le verbe *noter*, il fournit souvent une information supplémentaire, sur laquelle l'auteur tient à attirer l'attention du lecteur [14] :

[14] Nous n'avons pas reproduit ici les effets d'une taxation plus forte des revenus financiers des ménages, parce qu'ils sont un peu complexes à analyser. **On peut noter cependant qu'**à cause d'un effet Laffer, les conséquences budgétaires ex post de cette mesure sont faibles et incertains. (eco.xml-s7289)

Pour la construction avec un complément nominal, le verbe *observer* est le plus productif [15] :

[15] Quelle que soit la condition, **nous observons une augmentation** des durées de tenue lorsque l'objet familier est présenté à nouveau lors de la phase test de reconnaissance. (psycho.xml-s12966)

Lorsque ce dernier accepte un nom de processus, cette construction peut alterner avec la construction à complétive : *nous observons que les durées de tenue augmentent lorsque...* Quant aux deux autres verbes, cette construction avec un complément nominal est beaucoup moins fréquente, mais elle est tout de même présente.

À cela s'ajoute la construction en *comme* qui semble être privilégiée par le verbe *observer*. Comme le souligne Grossmann et Tutin (2010a, 2010b), cette construction marque une « fonction métatextuelle et/ou évidentielle », soit pour faciliter le guidage de lecture en ajoutant une preuve telle que les données figurant sur un graphique (*comme on l'observe sur le graphique*) [16], soit pour renforcer son argument à

travers les propos d'un autre chercheur (*comme l'observent Morris et Jones (1990)*) [17] :

[16] **Comme on l'observe sur le graphique**, ce sont donc principalement les localités au nord du corridor qui ont connu une augmentation importante de la population. (geo.xml-s10734)

[17] **Comme l'observent Morris et Jones (1990)**, le renouvellement permanent du contenu de la boucle phonologique est un aspect singulier de la tâche qui requiert une coordination rigoureuse de la part du centre exécutif. (psycho.xml-s13522)

Dans les deux exemples ci-dessus, le verbe utilisé avec une valeur évidentielle est associé à une fonction argumentative. En dehors de ces constructions syntaxiques principales, il nous semble que les verbes de constat sont facilement utilisés dans des contextes impersonnels, en particulier le verbe *noter* (*il est à noter que, il convient de noter que, il est intéressant de noter que*) [18].

[18] Enfin, **il est intéressant de noter que**, même parmi les enfants qui réussissent les deux épreuves, aucun ne présente un score de réussite plus élevé en production qu'en révision. (psycho.xml-s11430)

Ainsi se confirme le fait que ce dernier se caractérise par une valeur d'ajout ou de commentaire, servant à renforcer l'argument précédent (Grossmann, 2014b). En outre, le passif et le passif pronominal sont relevés pour le verbe *observer*. Le verbe *constater* se distingue des autres verbes par une fréquence assez élevée dans la construction *force est de constater que*, notamment utilisée pour aborder un résultat de recherche ou un argument souvent largement reconnu et incontestable [19].

[19] Si on a longtemps cru que l'appartenance à un groupe dévalorisé par la société se traduisait par des effets négatifs sur l'estime de soi, **force est de constater que** ce groupe peut revêtir des vertus autoprotectrices... (psycho.xml-s12776)

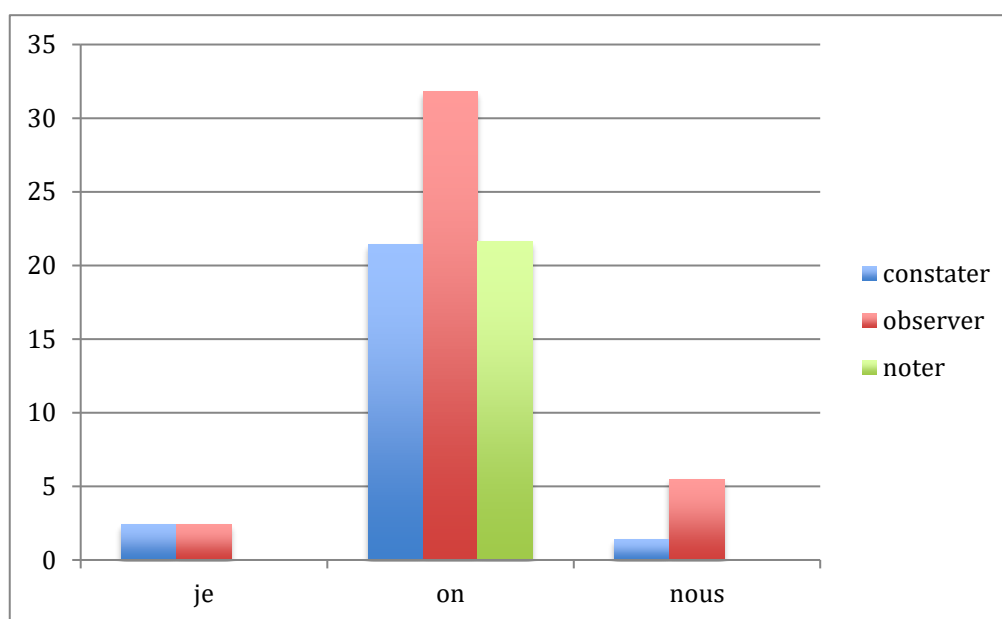
Les trois verbes de constat présentent des constructions syntaxiques assez variées et intéressantes sur le plan didactique, que nous récapitulons dans le tableau 4.2.

Verbes	Constructions syntaxiques récurrentes
observer	<i>qch. s'observe</i> (42/1428), <i>qch. est observé</i> (26/1428), <i>si l'on observe</i> (21/1428), <i>permettre d'observer</i> (19/1428), <i>comme l'observe qn.</i> (17/1428),
constater	<i>force est de constater que</i> (62/619), <i>permettre de constater</i> (22/619), <i>il est intéressant/frappant/paradoxal de constater que</i> (20/619)
noter	<i>il faut noter que</i> (63/250), <i>il est à noter que</i> (42/250), <i>il est intéressant/important de noter que</i> (39/250), <i>il convient de noter que</i> (8/250)

**Tableau 4.2 : Récapitulatif des constructions syntaxiques récurrentes des verbes de constat**

### b. Pronoms sujets mettant en jeu l'auteur

Comme nous l'avons souligné plus haut, les verbes de constat mobilisent un sujet humain et celui-ci renvoie souvent à l'auteur. En revanche, si l'on considère les pronoms sujets mettant en jeu l'auteur (*je*, *on* et *nous*), on constate que c'est le pronom *on* qui est le plus fréquemment employé par les trois verbes, de manière assez nette, comme illustré dans la figure 4-4 :



**Figure 4-4 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de constat)**

Il s'agit de l'emploi inclusif du pronom *on* qui remplace l'auteur et le lecteur (Fløttum et al., 2006). La prédominance de cet emploi rejoint le constat de Grossmann et Tutin (2010b), selon lequel, l'utilisation du *on* dans le constat scientifique est une marque de « co-constat » où l'auteur interagit avec le lecteur, permettant par conséquent « une forme de négociation dans l'argumentation, renforcée par l'emploi

d'atténuateurs » (*ibid.*, p. 7). Ici, l'emploi d'atténuateurs peut se rapporter au modalisateur *pouvoir* [20] :

[20] **On peut noter que**, même dans *Surveiller et punir*, Foucault insiste peu, par rapport à Guy Vincent, sur le rôle de l'organisation du savoir en "disciplines" ou de l'apparition de l'"apprentissage par principes" dans l'invention de l'école comme forme disciplinaire. (scienceseducation.xml-s10219)

Comme le montre l'exemple, le verbe *noter* associé au pronom *on* remplit la fonction d'inviter le lecteur comme témoin dans l'observation de la preuve, mettant en évidence la validité de son discours. Par ailleurs, les pronoms *nous* et *je* sont beaucoup moins fréquents, mais ils sont présents pour les verbes *observer* et *constater*. Par exemple, le verbe *observer* peut s'associer au pronom *nous* renvoyant à l'auteur (singulier ou collectif) [21].

[21] Cet effet, assez fréquemment décrit, s'observe à la fois lors de tâches de production, de reproduction de durées et lors de tâches d'estimation verbale (Wearden, 2003). **Nous l'avons également observé** dans notre tâche de référence... (psycho.xml-s11908)

En revanche, les pronoms *je* et *nous* sont presque absents pour le verbe *noter*, ce qui pourrait s'expliquer par l'importance de sa valeur de commentaire qui consiste à attirer l'attention du lecteur, et de ce fait, c'est le *on* incluant l'auteur et le lecteur qui est prédominant.

### c. Arguments nominaux

Pour les verbes de constat, les sujets nominaux récurrents se rapportent à l'auteur ou à un chercheur cité. Nous nous intéressons aux objets noms qui sont assez variés, en particulier pour le verbe *observer*, comme illustré dans la figure 4-5 :



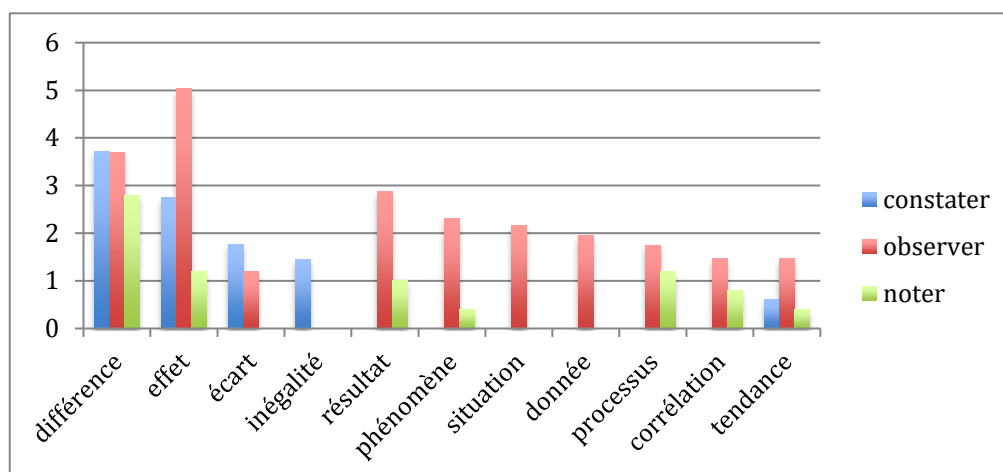


Figure 4-5 : Les noms objets des verbes de constat (en %)

Le verbe *observer* peut prendre comme objet des noms qui renvoient à un objet observable (*donnée*), à une relation (*différence*, *effet*, *écart*, *corrélation*, *résultat*), à un processus (*augmentation*, *processus*) ou à un cas (*phénomène*, *situation*) [22] :

[22] Ces **données ont été observées** chez les quatre groupes d'enfants de la première à la dernière année d'apprentissage de la lecture. (psycho.xml-s13934)

Le verbe *noter* se rapproche du verbe *observer* en ce qu'il accepte un nom de processus (*processus*), un nom qui marque une relation (*différence*, *effet*, *corrélation*, *résultat*) ou un nom relevant d'un cas scientifique (*phénomène*, *situation*). Par rapport à ces deux verbes, le verbe *constater* se caractérise par « une dimension plus abstraite », moins utilisé dans un emploi empirique (Grossmann, 2014b, p. 97), ce qui explique pourquoi il choisit rarement un nom concret comme *donnée*, *corpus*, etc.

Enfin, il nous semble que les trois verbes sont étroitement liés au co-constat. L'association entre les verbes de constat et le pronom *on* inclusif a la fonction principale d'inviter le lecteur dans la découverte de la preuve, en les associant à l'argumentation (Fløttum et al., 2006 ; Grossmann & Tutin, 2010b). Par contre, ils apparaissent différents tant au niveau syntaxique que sémantique. Contrairement au verbe *observer* « davantage mobilisé pour décrire la phrase empirique de la recherche » (Grossmann, 2014b, p. 99), *constater* et *noter* ont une dimension plus abstraite, revêtant une valeur argumentative plus importante. Si le verbe *constater*, davantage employé dans la construction en complétive *que*, introduit un résultat, le verbe *noter*, fréquemment utilisé dans les contextes impersonnels, introduit le plus

souvent une information complémentaire.

Nous présentons les patrons lexico-syntaxiques de ces trois verbes dans le tableau 4.3.

observer	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] (pouvoir) observe [complétive <i>que</i>] Ex : <b>Nous avons également observé que</b> le lestage est précédé d'une flexion du coude droit dès 20 mois. (psycho.xml-s10626)</li> <li>2. [[humain=auteur, chercheur]] (pouvoir) observe [[objet observable relation processus cas]] Ex : <b>Dans le cas de la révision d'un accord verbal erroné, on observe un phénomène original, inexistant en révision d'accord nominal erroné.</b> (psycho.xml-s11465)</li> <li>3. [[objet observable relation processus cas]] s'observe Ex : <b>Le rapport symétrique s'observe également dans le cas de l'emploi en interposition.</b> (ling.xml-s11193)</li> <li>4. Comme l'observe [[humain=auteur, chercheur]] Ex : <b>Comme l'observent Morris et Jones (1990), le renouvellement permanent du contenu de la boucle phonologique est un aspect singulier de la tâche...</b> (psycho.xml-s13522)</li> <li>5. [[activité scientifique support]] permet d'observer [[objet observable relation processus cas]] Ex : <b>Le tableau 3 permet d'observer cette mutation à une échelle plus petite.</b> (geo.xml-s9200)</li> </ol> <p>Lexset [[objet observable relation processus cas]] : <i>effet, différence, résultat, phénomène, situation, donnée, processus, tendance</i> Sens : Quelqu'un prend connaissance de la réalité sur un objet d'étude, d'une relation, d'un phénomène au moyen d'observations scientifiques.</p>
constater	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] (pouvoir) constate [complétive] Ex : <b>On peut ainsi constater que</b> l'effet du positionnement partisan est statistiquement significatif sur le choix des normes économiques. (sciencespolitiques.xml-s1119)</li> <li>2. [[humain=auteur, chercheur]] (pouvoir) constate [[relation propriété]] Ex : <b>Mais nous ne constatons pas de différence concernant le manque de ressources.</b> (psycho.xml-s15876)</li> <li>3. [[activité scientifique support]] permet de constater [complétive <i>que</i>] Ex : <b>Entretiens comme observation ont permis de constater que cette attitude reste rarement retrouvée dans le secteur étudié.</b> (scinfo.xml-s11902)</li> </ol> <p>Lexset [[relation propriété]] : <i>effet, différence, écart, inégalité</i> Sens : Quelqu'un prend connaissance de la réalité sur la relation ou les caractéristiques de quelque chose au moyen d'observation scientifique.</p>
noter	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] (pouvoir) note (adverbe discursif) [complétive <i>que</i>] Ex : <b>On peut noter que</b> le facteur d'impact de certaines revues en libre accès s'envole. (scinfo.xml-s13373)</li> <li>2. [[humain=auteur, chercheur]] (pouvoir) note [[relation processus cas=information supplémentaire]] Ex : <b>De même, on peut noter une différence d'ordre morpho-syntaxique entre les prépositions...</b> (ling.xml-s1967)</li> </ol> <p>Lexset [[relation processus cas=information supplémentaire]] : <i>différence, résultat, processus, corrélation, phénomène, tendance</i> Sens : Quelqu'un introduit une information supplémentaire et invite le lecteur à y prêter l'attention.</p>
Valeurs	Valeur argumentative, valeur évidentielle

sémantiques	
Routines	<i>si l'on observe</i> (21/1428) <i>force est de constater que</i> (62/619) <i>il est intéressant/frappant/paradoxal de constater que</i> (20/619) <i>il faut noter que</i> (63/250) <i>il est à noter que</i> (42/250) <i>il est intéressant/important de noter que</i> (39/250) <i>il convient de noter que</i> (8/250)
Commentaires	le pronom <i>on</i> est très fréquent. Le verbe <i>noter</i> est fréquemment associés aux adverbes discursifs tels que <i>cependant, également, par ailleurs, toutefois, tout d'abord</i> .

**Tableau 4.3 : Patrons lexico-syntaxiques des verbes de constat**

### 4.3.3. Verbes de description

La description scientifique joue un rôle important pour faire connaître au lecteur les données collectées, la notion ou la méthode adoptée. L'activité de description scientifique consiste à représenter et mettre en évidence les traits apparents d'un objet d'étude à l'aide d'un moyen, qu'il s'agisse soit d'un support graphique soit d'une formulation. Dans le LVF, les verbes de description relevés appartiennent à la classe générique « communication » et partagent le sens de base 'décrire qch à qn au moyen de'. Plus précisément, il s'agit de 'représenter en détail par écrit ou oralement, certains traits apparents' (d'après le TLFi). Nous présentons trois verbes représentatifs de cette sous-classe *décrire* (1127), *représenter* (1635), *détailler* (133).

#### a. Constructions syntaxiques et valeurs sémantiques

Ces verbes de description peuvent apparaître dans la construction transitive à sujet humain et à objet inanimé et peuvent être suivis par un circonstant de manière, comme le montre l'exemple [23] :

[23] **On peut représenter** les principaux types de théories pragmatiques par les parcours suivants. (ling.xml-s11633)

Cependant, nous avons remarqué que cette construction syntaxique est beaucoup moins fréquente que le passif dans notre corpus d'analyse. Lorsque le verbe *décrire* figure dans des structures au passif, il appelle fréquemment un adverbe de temps (*précédemment*) ou d'espace (*ici, ci-dessous*) permettant la navigation textuelle, avec par conséquent une valeur métatextuelle [24] :

- [24] Ce mécanisme **peut être décrit ici** : on sélectionne le plus souvent les propos que l'on tient en fonction de ce que l'on croit savoir des personnes à qui l'on s'adresse. (sciencespolitiques.xml-s10465)

Des adverbes de manière (*précisément, brièvement*) ont également été relevés [25] :

- [25] Cette précaution prise, on peut maintenant tenter, à la lumière des réponses aux questions de l'enquête mensuelle, de **décrire plus précisément** les fluctuations conjoncturelles que l'industrie française a traversées. (eco.xml-s2207)

Les verbes *représenter* et *détailler* sont en revanche fort utilisés avec un localisateur renvoyant à un élément péri-textuel, comme dans les exemples [26] et [27] :

- [26] Des phénomènes, des processus et des structures, forcément complexes, **sont ainsi représentés dans une figure** synthétique qui peut avoir valeur de paradigme reconnu. (geo.xml-s10905)
- [27] La résolution du problème de l'employeur **est détaillée en annexe B**. En menant un raisonnement symétrique à celui du paragraphe précédent, on vérifie tout d'abord que les entreprises structurellement déficitaires n'ont jamais intérêt à embaucher sur CDI. (eco.xml-s5367)

En effet, le verbe *décrire* se distingue de deux autres verbes par sa valeur métatextuelle qui renvoie souvent au processus d'écriture comme les verbes *présenter* et *montrer*. Cet emploi concerne notamment l'introduction et la conclusion, en particulier lorsque le verbe est employé au passé ou au futur, comme le montre l'exemple [28] :

- [28] **Nous avons donc décrit**, à travers l'analyse des dossiers disciplinaires, cinq processus qui apportent des éléments de compréhension de l'écart... (sciencespolitiques.xml-s14930)

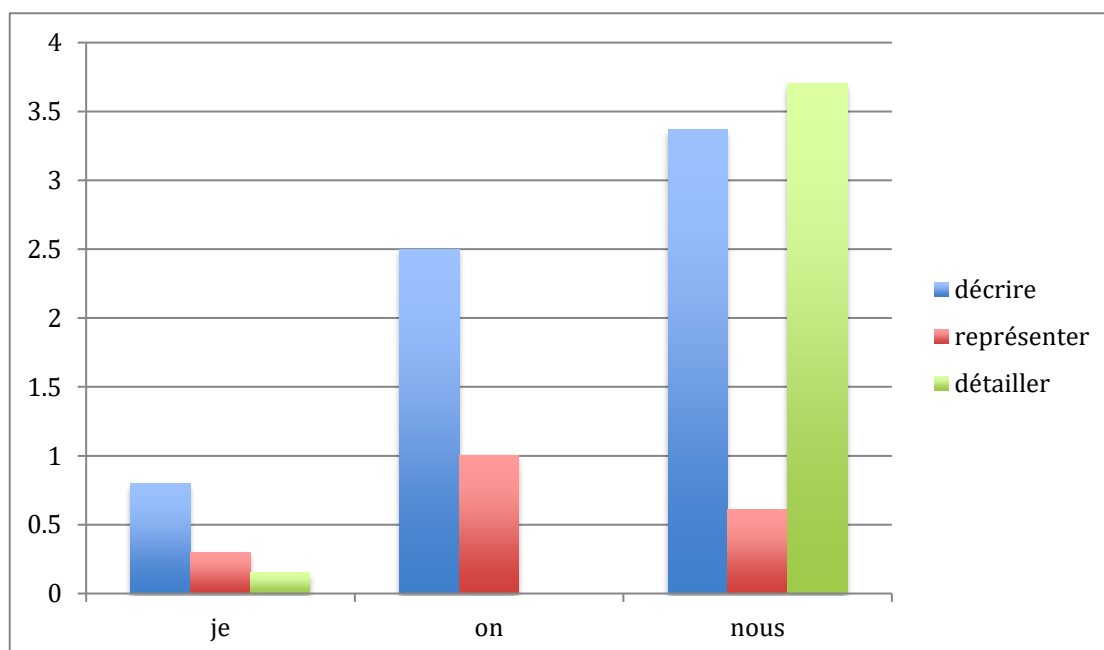
Un autre point différent à souligner est lié aux différents moyens de description mobilisés par ces trois verbes : le verbe *décrire* met en évidence l'activité de description au moyen de la formulation (*décrire brièvement*) ; le verbe *représenter*

rend perceptibles les traits à l'aide d'un schéma ou un graphique (*représenter dans la figure suivante*) ; enfin, le verbe *détailler* fait apparaître les moindres éléments liés aux caractéristiques (*détailler en annexe*). Ces différences sémantiques expliquent ainsi leur préférence de circonstants de manière.

Nous avons également relevé la présence de la construction relative pour *décrire* et *représenter* qui permet de mettre en avant l'objet à décrire comme dans *la méthode que nous décrivons* (*décrire* : 80/1127).

### b. Pronoms sujets mettant en jeu l'auteur

Au niveau du sujet, on constate que les pronoms mettant en jeu l'auteur sont peu fréquents, ce qui peut s'expliquer par la prédominance du passif pour les trois verbes, montrant ainsi que la présence de l'auteur est implicite dans l'activité de description. Les résultats obtenus sont illustrés dans la figure 4-6 :



**Figure 4-6 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de description)**

Il apparaît que le pronom *nous*, renvoyant à l'auteur (singulier ou collectif) est relativement plus fréquent [29] et [30] :

[29] Afin de discuter les théories de la convergence et des digital natives, **nous allons à présent décrire trois cas** de pratiques de jeunes observées sur les médias sociaux.

(scinfo.xml-s10714)

- [30] Enfin, **nous détaillons l'ensemble des extensions** possibles à partir de ces premiers travaux économétriques. (eco.xml-s9607)

Dans les exemples ci-dessus, le *nous* renvoie au processus d'écriture, marquant ainsi une valeur métatextuelle.

Le pronom *on* est moins présent ou même absent. Par ailleurs, ses valeurs variées sont difficiles à interpréter. Comparons les exemples [31] et [32] :

- [31] Dans cette figure, **on représente aussi les liens** thématiques concernant les fonctionnalités lieu - destination et localisable - déplaçable. (ling.xml-s8010)

- [32] Je crois que tout cela n'a pas été vu en grande partie parce que l'**on représente généralement l'organisation** en sections selon la figure 3 qui privilégie la filiation par les femmes (tout comme la figure 4 privilégie la filiation par les hommes). (anthropologie.xml-s15791)

Il nous semble que dans l'exemple [31], le *on* représente l'auteur (singulier ou collectif). Dans l'exemple [33], le *on* peut renvoyer aux autres chercheurs dont la méthode est remise en cause par l'auteur. Ces emplois diversifiés témoignent de la complexité du pronom *on* dans l'écrit scientifique.

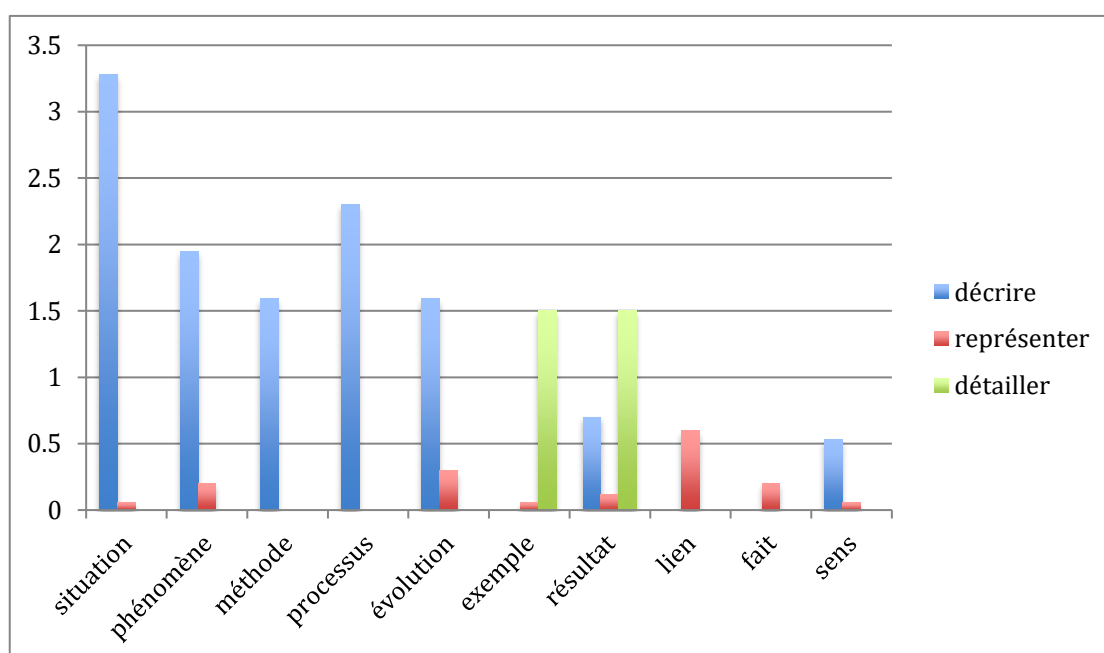
### c. Arguments nominaux

À la place du sujet, des cas d'emploi métonymique sont relevés pour les verbes *décrire* et *représenter* où l'auteur s'efface derrière la partie textuelle (paragraphe, partie du texte : *la section décrit X = nous décrivons X dans la section*) ou le moyen de description (graphique, figure : *cette figure représente X = nous représentons X dans la figure*), comme en témoigne l'exemple [33] :

- [33] **La Figure 2 représente l'ensemble de la démarche** mise au point pour concevoir et développer le système d'assistance VIVRE2. (psycho.xml-s5076)

Ici, les sujets métonymiques permettent de produire un effet objectivant dans la description scientifique. Concernant les objets, on rencontre des noms assez variés en

fonction objet, mais on peut essayer d'en tirer des généralisations par rapport à ce que l'on décrit. Tout d'abord, le verbe *décrire* peut sélectionner un objet scientifique (*méthode, phénomène, situation, sens*), un processus (*évolution, processus*) ou une relation (*résultat, différence*). Ensuite, le verbe *représenter* appelle plus fréquemment que les autres un nom abstrait (*lien, fait*) qui nécessite d'être décrit à l'aide d'un élément graphique ou visuel. Enfin, le verbe *détailler* se rapproche du verbe *décrire* dans la mesure où il préfère des noms renvoyant à l'objet d'étude (*exemple, modèle*) ou au rapport entre deux quantités (*résultat*), ces noms étant souvent au pluriel. Nous présentons les objets nominaux les plus fréquents dans la figure 4-7 :



**Figure 4-7 : Les noms objets des verbes de description (en %)**

L'essentiel dans la description scientifique est de faire connaître un objet scientifique en mettant en relief ses traits. Il ne s'agit pas de présenter l'agent qui a fait la description. Les objets décrits que sont les méthodes, les processus ou les résultats sont les acteurs du premier plan, comme en témoigne l'indication fréquente du passif et de la construction relative. Nous avons aussi montré comment les trois verbes mobilisent différents moyens de description en fonction de leur propre sémantisme. Ces propriétés sémantiques et syntaxiques seront donc intégrées dans les patrons verbaux récapitulés dans le tableau 4.4 :

décrire	1. [[humain=auteur, chercheur]] décrit (adverbe de temps espace manière) [[objet scientifique processus relation]] Ex : <i>C'est la construction de ce patron que nous allons décrire à présent.</i> (ling.xml-s13372) Lexset [[objet scientifique processus relation]] : <i>situation, phénomène, méthode, processus, évolution, différence</i> Sens : Quelqu'un met en évidence les traits apparents d'un objet d'étude, d'un processus ou d'une relation au moyen de la formulation.
détailler	2. [[humain=auteur, chercheur]] détaille (adverbe de manière) [[objet scientifique]] (dans [[élément péri-textuel]]) Ex : <i>Dans ces articles, les rédacteurs détaillent de manière distanciée les résultats du sondage et collectent diverses réactions.</i> (scinfo.xml-s10357) Lexset [[objet scientifique]] : <i>exemple, modèle, résultat</i> Sens : Quelqu'un fait apparaître (en annexe) les moindres éléments liés aux caractéristiques d'un objet d'étude.
représenter	3. [[humain=auteur, chercheur]] représente (adverbe de manière) [[cas relation]] (dans [[élément péri-textuel=visuel]]) Ex : <i>Dans cette figure, on représente aussi les liens thématiques concernant les fonctionnalités lieu - destination et localisable - déplaçable.</i> (ling.xml-s8010) Lexset [[cas relation]] : <i>phénomène, évolution, lien, fait</i> Sens : Quelqu'un rend perceptibles les traits liés à un cas ou une relation à l'aide d'un schéma ou un graphique.
Valeurs sémantiques	Valeur descriptive, valeur métatextuelle
Routines	(la méthode) que nous décrivons (80/1127), pour décrire (70/1127)/représenter (40/1635)
Commentaires	L'emploi métonymique est possible au niveau du sujet ( <i>section, figure</i> ). Le passif et la construction relative sont très fréquents, le pronom <i>nous</i> est fréquent. Le verbe <i>décrire</i> appelle fréquemment un adverbe de temps ( <i>précédemment</i> ), d'espace ( <i>ici, ci-dessous</i> ), ou de manière ( <i>précisément, brièvement</i> )

**Tableau 4.4 : Patrons lexico-syntactiques des verbes de description**

#### 4.3.4. Verbes d'examen

De même que les verbes de constat, les verbes d'examen relevés sont interprétés sémantiquement dans le LVF comme des verbes psychologiques et partagent le sens de base 'manifester telle activité réflexive sur qn/qc'. Plus précisément, ces verbes signifient 'étudier en détail, faire l'analyse de quelque chose pour en avoir une connaissance approfondie, juger de sa valeur' (d'après le TLFi). Dans la recherche scientifique, ces verbes sont notamment utilisés pour aborder et remettre en question un objet de recherche afin de mener des réflexions autour de cet objet, pour se former ainsi son propre point de vue. Ils permettent donc d'établir le lien entre le scripteur (rôle de chercheur) et son objet d'étude observable (*données*) ou abstrait (*approche, caractéristique, relation*). Nous proposons d'observer quatre verbes représentatifs



appartenant à cette sous-classe : *examiner* (498), *étudier* (971), *analyser* (955), *explorer* (212).

### a. Constructions syntaxiques et valeurs sémantiques

Pour les verbes étudiés, la structure privilégiée est une construction avec complément nominal [34] :

- [34] Dans cet article, **nous examinons l'influence** d'un évènement dans les pratiques communicationnelles développées au sein d'un territoire productif. (scinfo.xml-s12128)

Il est intéressant de constater que ces verbes sont fréquemment associés aux adverbes d'énumération (*d'abord, ensuite*) ou aux déictiques de lieu (*ici, plus haut*). Ici, les verbes remplissent une fonction métatextuelle renforcée par l'indication du temps du passé ou du futur, comme le montrent les exemples [35] et [36] :

- [35] **J'examinerai d'abord** l'arrière-plan institutionnel de la croissance de la scolarisation, c'est-à-dire en particulier les conséquences de deux décisions qui n'ont pas suscité de controverses aiguës et peuvent paraître mineures. Je m'attacherai ensuite à... (histoire.xml-s5599)

- [36] **Nous étudions ici** des énoncés figuraux basés sur l'expression de références plurielles (syntagmes nominaux définis pluriel, noms collectifs), dans un corpus journalistique portant sur l'actualité. (ling.xml-s2766)

Cette sous-classe se rapproche de celle de description dans la mesure où les quatre verbes étudiés présentent aussi une préférence pour la construction au passif [37] et la construction relative [38] permettant de mettre en avant l'objet d'étude :

- [37] **Les effets** de la discrimination du troisième degré sur les prix et le bien-être **ont été** le plus souvent **analysés** dans un cadre monopolistique (ROBINSON [1933], VARIAN [1985, 1988]). (eco.xml-s11423)

- [38] **Les situations que nous étudions** dans nos recherches sont des situations de classe que nous qualifions d'ordinaires, au sens où elles ne sont pas issues d'une ingénierie didactique. (scienceseducation.xml-s11429)

Des contextes impersonnels à fonction déontique (*s'agir, importer, convenir*) ont également été relevés [39] :

[39] Maintenant que les conséquences des phénomènes d'agrégation sociale et/ou académique sur les performances des élèves sont documentées, **il importe d'en étudier les effets** sur l'équité des systèmes éducatifs. (scienceseducation.xml-s2358)

L'observation du corpus nous permet de constater que ces verbes sont utilisés dans des constructions qui ont des fonctions différentes :

1. introduire et mettre en avant l'objet d'étude [40] :

[40] **Il s'agit à présent d'examiner** comment et dans quelle mesure les services installés sur les archives ouvertes répondent déjà à ces objectifs. (scinfo.xml-s12310)

2. introduire l'intérêt de l'étude en soulignant l'originalité ou la motivation [41] :

[41] Notre approche est celle d'une géographie culturelle appliquée à l'animal, domaine encore **très peu exploré** par les géographes français. (geo.xml-s1883)

3. exprimer l'objectif de l'étude [42] :

[42] **Cet article se propose d'analyser** le rôle des outils spirituels et politiques dans l'élaboration de cette communauté politique, et de mesurer leur importance de nos jours après plusieurs situations de crise. (anthropologie.xml-s467)

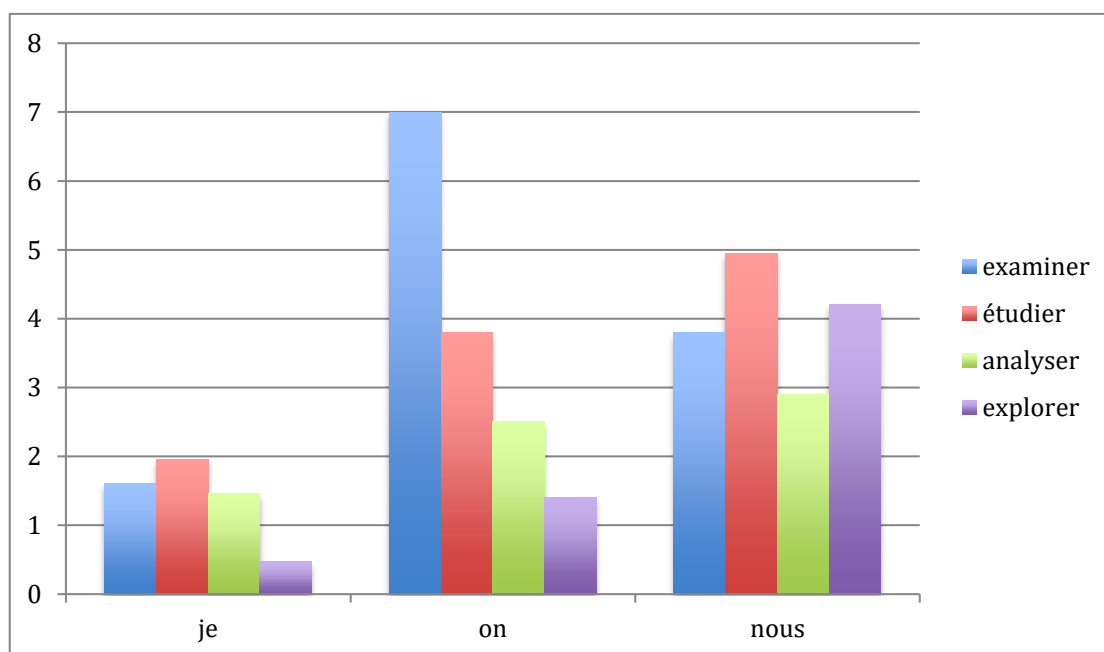
Afin de mettre en relation les constructions syntaxiques récurrentes et leurs fonctions correspondantes, nous proposons ci-dessous un tableau récapitulatif :

Fonctions	Constructions syntaxiques récurrentes
Introduire et mettre en avant l'objet d'étude	<ol style="list-style-type: none"> <li>1. construction relative : <i>que nous étudions</i> (66/971), <i>que nous analysons</i> (41/955)</li> <li>2. passif : <i>être analysé</i> (84/955), <i>être étudié</i> (80/971)</li> <li>3. gérondif : <i>en analysant</i> (29/955) <i>en examinant</i> (15/498), <i>en étudiant</i> (43/971)</li> <li>4. <i>il s'agit d'étudier</i> (10/971)/<i>examiner</i> (6/498)/<i>analyser</i> (6/955)</li> <li>5. <i>si l'on examine</i> (9/498)/<i>analyse</i> (10/955)</li> </ol>
Introduire l'intérêt de l'étude	<ol style="list-style-type: none"> <li>1. <i>être peu étudié</i> (23/971)/<i>exploré</i> (20/212)</li> <li>2. <i>il convient/importe d'étudier</i> (10/971)/<i>examiner</i> (8/498)</li> </ol>
Exprimer l'objectif de l'étude	<ol style="list-style-type: none"> <li>1. <i>pour étudier</i> (82/971)/<i>examiner</i> (8/498)/<i>analyser</i> (66/955) /<i>explorer</i> (18/212)</li> <li>2. <i>(se) proposer d'étudier</i> (8/971)/<i>examiner</i> (6/498)/<i>analyser</i></li> </ol>

**Tableau 4.5 : Récapitulatif des constructions syntaxiques récurrentes des verbes d'examen et leurs fonctions correspondantes**

### b. Pronoms sujets mettant en jeu l'auteur

En ce qui concerne les pronoms sujets, nous avons relevé une fréquence relativement élevée pour les pronoms *nous* et *on*. Contrairement à d'autres verbes étudiés précédemment qui montrent une préférence marquée pour un seul pronom, ces quatre verbes peuvent sélectionner soit le pronom *nous* et soit le pronom *on* en fonction du contexte. La figure 4-8 donne un aperçu de leurs fréquences relatives dans le corpus :



**Figure 4-8 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes d'examen)**

Quand le pronom *on* est employé, il semble que l'auteur souhaite impliquer le lecteur dans l'activité d'examen pour en faire un témoin, et ceci ne sera pas le cas pour le verbe *décrire* [43] :

[43] Nous cherchons tout juste à mettre en exergue quelques faits caractéristiques. **On examinera ainsi**, en revenant sur des travaux plus anciens (notamment Cadiot 1999 et Cadiot & Visetti 2001) quelques prépositions impliquées dans l'expression

grammaticale de l'espace. (ling.xml-s2570)

En outre, la présence de certaines constructions employant le pronom *on* et comportant les conjonctions *si/lorsque/quand* en apporte aussi la preuve [44] :

- [44] Pourtant, **si l'on examine** attentivement cette table de scolarité située dans l'annexe de l'ouvrage « Population » et l'enseignement (INED, 1970, p. 446-447) et partiellement reproduite ci-dessous (voir le tableau 1), on note le résultat contraire. (scienceseducation.xml-s3044)

Le pronom *nous* renvoie souvent à l'auteur (singulier ou collectif) et le verbe mobilise une valeur métatextuelle. Dans les exemples suivants, la construction comportant le *nous*+verbes d'examen a pour fonction principale d'introduire l'objectif de l'étude [45] ou de faire un rappel dans la conclusion [46] :

- [45] Dans le cadre du présent article, **nous voudrions explorer** ces aspects contradictoires des activités de recherche de reconnaissance sur les SRS. (scinfo.xml-s9126)
- [46] **Dans cet article, nous avons analysé** par quels mécanismes le niveau des allocations chômage peut affecter le taux de chômage de long terme dans le cadre d'un modèle de négociations salariales. (eco.xml-s1165)

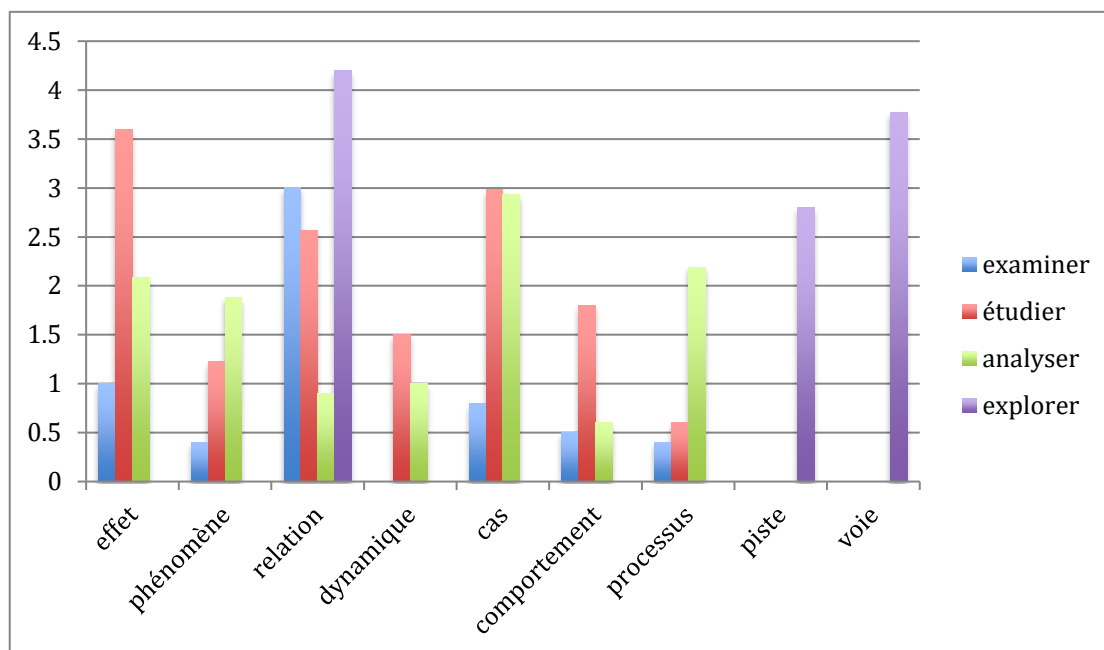
Enfin, il faut souligner que l'emploi du pronom *je* reste marginal, mais il n'est pas absent de notre corpus.

### c. Arguments nominaux

Des emplois métonymiques (*article, section*) ont été constatés, qui consistent en une autre façon de marquer la présence de l'auteur. Des associations comme *l'article examine* (23/498)/*étudie* (27/971) ou *la section étudie* (10/971)/*analyse* (8/955) sont ainsi naturelles. Ces mentions dépersonnalisées de l'auteur produisent un effet objectivant.

Quant aux noms objets, les quatre verbes sont productifs. Deux cas se détachent nettement : d'un côté, les verbes *examiner, étudier* et *analyser* peuvent sélectionner des noms relevant d'un processus (*processus, évolution, dynamique, comportement*), d'une relation (*effet, relation*) ou d'une propriété (*caractéristique*) ; d'un autre côté, le

verbe *explorer* se restreint à certains noms abstraits qui relèvent de l'orientation de la recherche (*relation*, *possibilité*, *piste* et *voie*). En effet, ce verbe mobilise le sens de découverte scientifique d'un sujet peu connu et caché, ce qui le distingue des autres verbes. Les co-occurrences les plus fréquentes sont illustrées dans la figure 4-9 :



**Figure 4-9 : Les noms objets des verbes d'examen (en %)**

Nous avons donc observé une variété de constructions syntaxiques et leurs fonctions liées à cette sous-classe. À l'instar des verbes de description, la mention de l'auteur est souvent implicite ou dépersonnalisée pour les verbes d'examen. L'importance dans l'examen ou dans l'analyse est de cibler et de mettre en avant l'objet d'étude en justifiant son intérêt. Cette caractéristique contribue à renforcer l'objectivité du fait et la validité du discours. Dans cette sous-classe, *examiner*, *étudier* et *analyser* semblent être sémantiquement assez proches, en particulier lorsqu'ils ont une valeur métatextuelle. Quant au verbe *explorer*, il est davantage lié au sémantisme de l'orientation de la recherche.

Placés de préférence au début ou à la fin du texte, les verbes d'examen ont un lien étroit avec l'objet d'étude et l'objectif d'étude et ils facilitent l'entrée en matière et présentent ainsi des fonctions importantes dans l'organisation textuelle.

Nous présentons les patrons lexico-syntaxiques de cette sous-classe dans le tableau 4.6 :

examiner	1. [[humain=auteur, chercheur]] examine (adverbe d'énumération espace) [[processus relation propriété]] Ex : <i>Plus récemment, Yoon et Suh (2003) ont examiné seulement trois liens avec la satisfaction.</i> (psycho.xml-s1594) Lexset : <i>effet, phénomène, relation, cas, processus, comportement</i> Sens : Quelqu'un aborde et fait l'analyse d'un objet d'étude ou d'un événement afin de déterminer ses propriétés.
étudier	2. [[humain=auteur, chercheur]] étudie (adverbe d'énumération espace) [[processus relation propriété]] Ex : <i>Dans ce papier, nous étudions les effets de l'assurance chômage et de l'assistance chômage dans un modèle de négociations salariales.</i> (eco.xml-s1017) Lexset : <i>effet, phénomène, relation, dynamique, cas, comportement, processus</i> Sens : Quelqu'un observe attentivement un objet d'étude ou un événement afin d'acquérir des connaissances.
analyser	3. [[humain=auteur, chercheur]] analyse (adverbe d'énumération espace) [[processus relation propriété]] Ex : <i>J'analyserai un cas particulier, celui d'une jeune thérapeute - possédée.</i> (anthropologie.xml-s11074) Lexset : <i>effet, phénomène, relation, dynamique, cas, comportement, processus</i> Sens : Quelqu'un fait l'analyse d'un objet d'étude ou d'un événement à l'aide des méthodes ou des outils spécifiques.
explorer	4. [[humain=auteur, chercheur]] explore [[orientation scientifique]] Ex : <i>Dans le cadre du présent article, nous voudrions explorer ces aspects contradictoires des activités de recherche de reconnaissance sur les SRS.</i> (scinfo.xml-s9126) Lexset [[orientation scientifique]] : <i>piste, voie, relation</i> Sens : Quelqu'un procède à l'étude systématique d'un fait scientifique, visant la découverte scientifique d'un sujet peu connu et caché.
Valeurs sémantiques	valeur métatextuelle, valeur introductive
Routines	<i>(la question) que nous étudions (66/971)</i> <i>(l'exemple) que nous analysons (41/955)</i> <i>il s'agit d'étudier (10/971)/examiner (6/498)/analyser (6/955)</i> <i>il convient/importe d'étudier (10/971)/examiner (8/498)</i> <i>être peu étudié (23/971)/exploré (20/212)</i> <i>pour étudier (82/971)/examiner (8/498)/analyser (66/955) /explorer (18/212)</i> <i>(se) proposer d'étudier (8/971)/examiner (6/498)/analyser (6/955)/explorer (4/212)</i> <i>si l'on examine (9/498)/analyse (10/955)</i>
Commentaires	L'emploi métonymique est possible au niveau du sujet ( <i>article, section</i> ). Le passif, la construction relative et le gérondif sont fréquents, le pronom <i>nous</i> et <i>on</i> sont fréquents. Les verbes <i>étudier, examiner</i> et <i>analyser</i> s'emploient souvent avec un adverbe d'énumération ( <i>d'abord, ensuite</i> ) ou un localisateur ( <i>ici</i> ).

Tableau 4.6 : Patrons lexico-syntaxiques des verbes de description

### 4.3.5. Verbes d'interprétation

L'activité d'interprétation consiste à mettre en place « une relation de preuve ou de justification » (Siepmann 2007, p. 129), caractéristique essentielle du discours scientifique dans lequel l'auteur doit établir un raisonnement rigoureux. Les trois

verbes représentatifs de cette sous-classe sont *interpréter* (487), *expliquer* (1635), *expliciter* (263). Ces verbes appartiennent à la classe générique « communication » dans le LVF et ont le même sens de base ‘montrer qc’. Ils admettent trois arguments : un objet d’interprétation, un agent et un destinataire, qui peut être présentés dans la formule suivante :

N [agent] +V+ N [objet d’interprétation] + à N [destinataire]

Notons par ailleurs que le troisième argument est souvent implicite. L’interprétation scientifique peut être considérée comme une étape d’analyse où l’auteur (agent) apporte au lecteur (destinataire) une explication ou une meilleure compréhension d’un fait exposé avant.

### a. Constructions syntaxiques et valeurs sémantiques

Pour les trois verbes étudiés, les structures privilégiées se montrent assez différentes : le verbe *expliquer* privilégie la construction avec une complétive (*nous expliquons que/comment/pourquoi*) [47] et avec un complément nominal (*nous expliquons cette différence*) alors que les verbes *interpréter* et *expliciter* préfèrent la construction avec un complément nominal (*nous interprétons/explicitons les conséquences*) [48]. Le verbe *interpréter* se distingue par le fait que son complément nominal est souvent suivi de la préposition *comme* qui introduit le sens attribué [49] :

[47] On essaiera donc d’**expliquer pourquoi** la forme des patrons ne correspond pas strictement aux structures linguistiques décrites. (linguistique\_422)

[48] C’est sur ce schéma que **nous expliciterons** plus en détail les différentes **étapes** qui conduisent à l’élaboration d’un patron complexe dont les performances en corpus sont satisfaisantes. (ling.xml-s13359)

[49] **On pourrait**, en ce sens, **interpréter** le double **caractère** équilibré et socialement indifférencié des opinions avancées **comme** l’expression d’une adhésion consensuelle à la liberté et à l’égalité. (sciencespolitiques.xml-s981)

De même que le verbe *interpréter*, *expliquer* peut s’accompagner de l’expression de cause introduite par la préposition *par* dont la fonction sémantique est similaire à

celle de *comme* [50] :

[50] Ne **peut-on expliquer** aussi **la faiblesse** des pratiques paysagères en forêt **par** une inadéquation structurelle entre les conceptions paysagistes existantes et l'approche territoriale de la forêt par ses gestionnaires ? (geo.xml-s10007)

Il est intéressant de constater que ces verbes sont fréquemment associés au modalisateur *pouvoir*. En effet, ces verbes sont utilisés pour fournir une interprétation ou une explication personnelle. La présence de ce marqueur permet d'atténuer l'assertion de l'auteur et d'accroître la crédibilité de son discours. Le lecteur peut accepter plus facilement le point de vue de l'auteur lorsque celui-ci est présenté de manière moins tranchée. Par ailleurs, le passif est très fréquent pour ces trois verbes [51] :

[51] Dans le répertoire de Minette, on trouve une Suvre dont le thème **pouvait être interprété comme** une allusion directe à la situation raciale dans la colonie. (histoire.xml-s2941)

Ici, la construction passive nous semble jouer une double fonction. D'une part, le fait d'éviter la mention de l'auteur permet de produire un effet objectivant, d'autre part, le passif contribue à présenter l'objet d'interprétation comme le thème de la phrase, ce qui a pour effet d'assurer une cohérence textuelle et d'attirer l'attention du lecteur sur la preuve et la justification, et non pas sur l'agent.

Par ailleurs, la construction causative comportant le verbe *permettre* est également relevée pour les trois verbes (en particulier le verbe *expliquer* qui marque 42 occurrences). Ici, sa fonction est d'établir une relation causale sur laquelle se fondent l'interprétation et l'explication de l'auteur. L'auteur fait comprendre son raisonnement rigoureux, comme en témoigne [52] :

[52] Le modèle précédent **permet d'explicitier** un des canaux de transmission de la politique monétaire et plus spécifiquement l'impact de cette dernière sur le taux des crédits distribués : la modification des taux directeurs affecte le coût des ressources. (eco.xml-s14765)

Nous avons également remarqué que le verbe *expliquer* est fréquemment utilisé



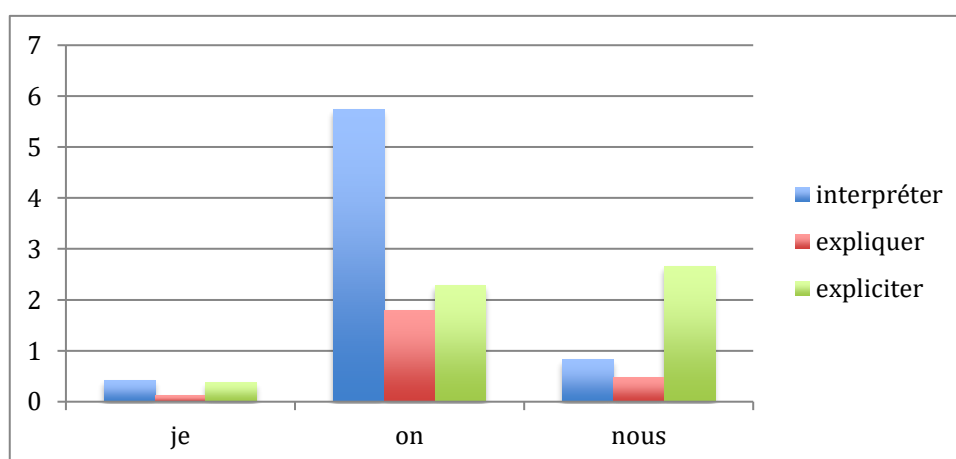
dans la construction comportant *comme* pour citer le propos d'un autre chercheur ou renvoyer à sa propre explication. Dans l'exemple [53], la construction rappelle le point de vue de l'auteur et contribue à la cohérence textuelle. Dans l'exemple [54], la construction aide à renforcer l'argument de l'auteur en le rapprochant d'un autre point de vue attesté dans le champ. Ici, le verbe prend une valeur métadiscursive, permettant d'« associer le lecteur à l'argumentation scientifique » (Grossmann & Tutin, 2010b, p. 3).

[53] **Comme nous l'avons expliqué dans la section 3.3**, le principe de l'apprentissage endogène des propriétés de complémentation s'effectue par le biais du repérage dans le texte de zones de rattachement non ambiguës. (ling.xml-s6738)

[54] L'ambiguïté du message est aussi illustrée par une forme en quelque sorte intermédiaire entre la ruine et le mémorial, assez répandue, **comme l'explique Jacques Sauvageot** : « Parmi ces commémorations, la pratique la plus courante consiste en l'idée de laisser.... » (Sauvageot, 1995). (geo.xml-s5805)

## b. Pronoms sujets mettant en jeu l'auteur

Afin d'observer la mention de l'auteur dans le processus de l'interprétation, nous proposons un aperçu des fréquences relatives de trois pronoms sujets étudiés dans le graphique 4-10 :



**Figure 4-10 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes d'interprétation)**

D'une manière générale, les trois pronoms sont peu fréquents pour cette

sous-classe, ce qui confirme que les verbes étudiés sont davantage utilisés au passif. Par contre, nous avons relevé une fréquence relativement élevée du pronom *on* pour le verbe *interpréter* qui mobilise souvent le modalisateur *pouvoir*. Dans l'exemple [55], l'association de *on* et *pouvoir* nous semble avoir une double valeur sémantique : d'une part, le *on* représente « une interprétation exclusive » (auteur+communauté de discours) (Tutin, 2010b, p. 36) dans laquelle l'auteur indique que sa manière d'interprétation n'est pas personnelle, mais partagée par la communauté scientifique. D'autre part, l'association *on* + *pouvoir* semble mettre en place une négociation avec le lecteur, « laissant la possibilité d'adhérer ou non » à son raisonnement (Hyland 2005b, cité par Grossmann & Tutin, 2010b, p. 3).

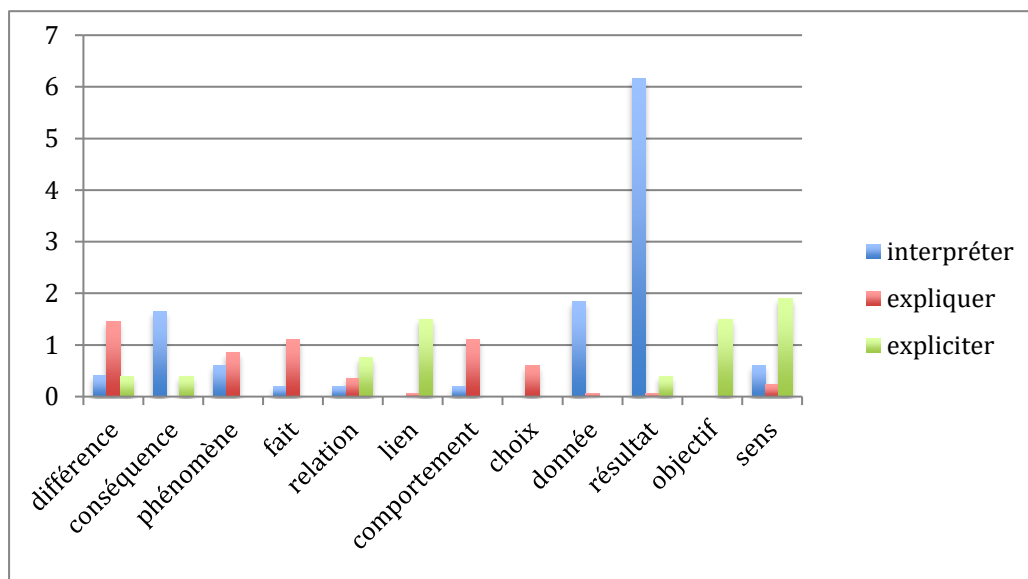
[55] Mais **on peut également interpréter** ce résultat en se référant, à la conception radicale de la dissonance cognitive, comme nous l'avons fait pour rendre compte a posteriori de l'effet surprenant observé dans la situation sans ultrason/centration de l'expérimentation 1 : le fait de se relire rend plus saillant les arguments trouvés. (psycho.xml-s10493)

### c. Arguments nominaux

Comme nous l'avons souligné, les verbes d'interprétation consistent à apporter une meilleure compréhension d'un phénomène ou d'un fait mentionné précédemment, qui peut être remplacé par le pronom neutre *le* dans la phrase, en particulier pour les verbes *l'expliquer* (59/1635)/*l'interpréter* (29/487) [56] :

[56] Ce n'est donc pas pour sa sophistication méthodologique, ou pour sa reconnaissance académique, que cet outil est revendiqué comme source d'inspiration pour la construction de la Bologna scorecard, **on peut plutôt l'expliquer** par une sorte de mode de cet outil (il est alors largement fait référence à la Lisbon scorecard dans la presse internationale [51] et dans les débats sur la stratégie de Lisbonne entre les élites des pays européens). (sciencespolitiques.xml-s12014)

Sémantiquement, les objets nominaux autorisés par les trois verbes étudiés sont assez différents. Les objets les plus fréquents sont illustrés dans la figure 4-11 :



**Figure 4-11 : Les noms objets des verbes d'interprétation (en %)**

En fait, le verbe *interpréter* est souvent associé aux noms renvoyant à la relation entre deux entités (*résultat, conséquence*) aux cas (*phénomène, fait*) ou aux données (*donnée*). Il s'agit plutôt des objets concrets et construits de l'analyse scientifique, à partir desquels, l'auteur donne son point de vue et formule une conclusion. Quant au verbe *expliquer*, il est difficile de l'associer aux objets concrets comme le nom *donnée*. Il apparaît que ce verbe a une dimension plus abstraite, sélectionnant par conséquent des noms relevant des types sémantiques [[processus]] (*comportement, choix*), [[cas]] (*fait, phénomène*) ou même [[relation]] (*différence, relation*). Enfin, le verbe *expliciter* se rapproche du verbe *expliquer* dans la mesure où il privilégie des noms renvoyant à une entité abstraite qui peut être la finalité (*objectif*), l'explicatif (*sens*) ou la relation (*différence, lien, conséquence*).

Les trois verbes étudiés ont montré des propriétés différentes tant au niveau sémantique que syntaxique. Les verbes *expliquer* et *expliciter* se rapprochent sémantiquement en ce qu'ils partagent des noms objets abstraits similaires alors que le verbe *interpréter* marque bien une prise de position de l'auteur, comme en témoigne le caractère récurrent de la construction *on peut interpréter*. Cependant, sur le plan argumentatif, ces verbes introduisent tous le raisonnement et la justification de l'auteur (*permet d'expliciter/expliquer, peut être interprété, on peut l'interpréter*) dont la responsabilité est souvent atténuée par le modalisateur *pouvoir*.

Interpréter	<p>1. [[humain=auteur, chercheur]] (pouvoir) interprète [[relation cas donnée]] (comme) Ex : <i>En effet, il interprète cette décision comme la rupture d'une relation de confiance...</i>(sciencespolitiques.xml-s10033)</p> <p>2. [[relation cas donnée]] (pouvoir) s'interprète (comme) Ex : <i>Cette condition s'interprète simplement...</i>(eco.xml-s16795)</p> <p>Lexset : <i>résultat, donnée, conséquence, différence</i> Sens : Quelqu'un donne un sens personnel à la relation entre deux entités, à un cas particulier ou aux données scientifiques afin d'en dégager un fait caché.</p>
Expliquer	<p>1. [[humain=auteur, chercheur]] (pouvoir) explique [[cas processus relation=conséquence]] (par [[élément=cause]]) Ex : <i>On ne peut donc expliquer le processus de périurbanisation uniquement à travers ses mécanismes physiques.</i> (geo.xml-s10296)</p> <p>2. [[humain=auteur, chercheur]] explique [complétive que/pourquoi/comment] Ex : <i>Shumway-Cook et Woollacott (1995) expliquent que le développement du contrôle postural repose sur celui d'au moins sept systèmes...</i>(psycho.xml-s10608)</p> <p>3. comme l'explique [[humain=auteur, chercheur]] Ex : <i>L'invocation des métaphores anatomique et psychanalytique devient courante, même si elle doit être nuancée, comme l'explique en particulier Françoise Choay [7].</i> (geo.xml-s5754)</p> <p>4. [[objet scientifique  activité scientifique=cause]] permet d'expliquer [[cas processus relation=conséquence]] Ex : <i>L'attention portée à cette première période permet d'expliquer les différentes facettes de cet activisme instrumental.</i> (sciencespolitiques.xml-s17238)</p> <p>Lexset : <i>différence, phénomène, fait, choix, comportement, relation</i> Sens : Quelqu'un fait connaître la raison de quelque chose ou d'un évènement, établissant ainsi un raisonnement rigoureux.</p>
Expliciter	<p>1. [[humain=auteur, chercheur]] (pouvoir) explicite [[finalité relation explicatif=complexe]] Ex : <i>Il explicite certaines caractéristiques d'Eurodac.</i> (sociologie.xml-s10443)</p> <p>Lexset : <i>lien, objectif, relation, caractéristique, sens</i> Sens : Quelqu'un rend plus clair et compréhensible ce qui a un sens vague et inconnu.</p>
Valeurs sémantiques	Valeur argumentative, valeur interprétative
Routines	<i>pour interpréter</i> (13/487)/ <i>expliquer</i> (112/1635)/ <i>expliciter</i> (7/263) <i>en interprétant</i> (4/487)/ <i>expliquant</i> (24/1635)/ <i>explicitant</i> (10/263)
Commentaires	Le passif est très fréquent, le gérondif est présent, le <i>on</i> est plus fréquent pour <i>interpréter</i> .

Tableau 4.7 : Patrons lexico-syntaxiques des verbes d'interprétation

### 4.3.6. Récapitulatif

Les verbes d'analyse étudiés peuvent être considérés comme des verbes de discours puisqu'ils ont des fonctions discursives importantes : guider le lecteur (la combinaison avec des expressions métatextuelles : *dans un premier temps*) ; impliquer le lecteur dans l'activité de recherche (l'association avec le pronom *on*

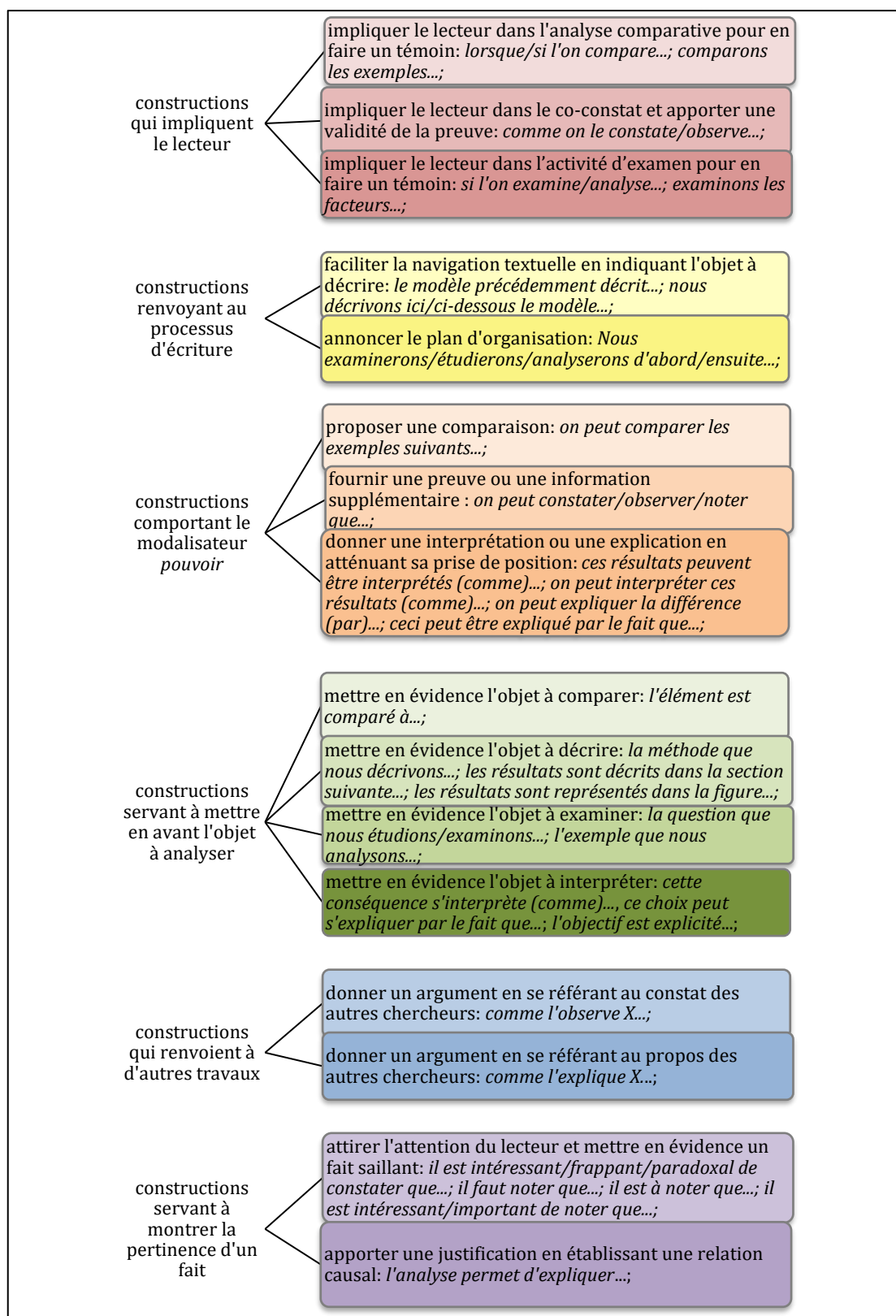
inclusif) ; mettre en relief les objets d'étude (l'apparition dans des constructions au passif, constructions relatives). Ces principales fonctions sont détaillées dans la figure 4-12.

De manière générale, à l'issue de l'analyse de cette sous-classe, plusieurs caractéristiques peuvent se dégager :

1. En dehors des emplois métatextuels, la mention de l'auteur est souvent implicite et dépersonnalisée (les sujets métonymiques pour les verbes de description et d'examen).
2. Les objets d'étude (les données, les méthodes ou les résultats) sont les acteurs du premier plan dans l'analyse scientifique, qu'ils soient le sujet de la phrase ou qu'ils apparaissent dans des contextes impersonnels (*il est à noter que*).
3. La prise en compte du lecteur est importante dans l'analyse scientifique. Le rôle du chercheur consiste à établir un lien entre ses objets d'étude et le lecteur : la co-comparaison des données, le co-constat des résultats, le co-examen des processus ou des méthodes, la négociation avec le lecteur dans l'interprétation et l'explication.
4. Les verbes d'analyse s'accompagnent souvent de l'expression de l'objectif (*pour étudier*), de l'intérêt (*il est intéressant de comparer*), et du moyen (*décrire plus précisément*).

Concernant les constructions verbales, comme nous l'avons expliqué plus haut, le nombre récurrent de constructions relatives (*la méthode que nous décrivons*), de passifs (*les données peuvent être interprétées comme*) témoigne du fait que dans l'analyse scientifique, la description des objets d'étude prime sur la présence de l'auteur. Par ailleurs, les différents types de constructions impersonnelles montrent que l'analyse scientifique consiste en un processus de recherche qui vise l'objectivité (Liddicoat, 2008 ; Fifielska, 2015) : les constructions avec verbes impersonnels introducteurs (*il faut noter que*) et les contextes impersonnels (*il est intéressant de constater*).

Nous présentons ci-dessous les constructions verbales identifiées de notre analyse, celles-ci sont organisées selon leur fonction rhétorique.



**Figure 4-12 : Les principales constructions verbales des verbes d'analyse et leurs fonctions**

#### 4.4. Les verbes de positionnement

Le terme de « positionnement » est utilisé ici comme générique qui permet de regrouper des verbes qui « indiquent un engagement fort de l’auteur, qu’il s’agisse de verbes d’opinion ou d’évaluation, ou de verbes indiquant un apport singulier de l’auteur dans la démonstration ou la découverte scientifique » (Tutin, 2010b, p. 24). En fait, la notion de positionnement est déjà abordée dans de nombreuses études qui s’accordent à confirmer qu’elle est « le résultat de procédés d’effacement énonciatif et de procédés de marquage de l’attitude ou de la subjectivité du locuteur » (Boch & Rinck, 2010, p.8). Ces études ont mis en avant les marqueurs linguistiques du positionnement, tels que les pronoms personnels (Fløttum, 2004 ; Reutner, 2010), les adjectifs évaluatifs (Tutin, 2010b), les verbes de démarcation (Chavez, 2008). Si l’on admet que l’expression du positionnement peut être véhiculée par diverses séquences lexicales, c’est à la suite de ces travaux que nous souhaitons étudier le positionnement à travers les patrons verbaux, éléments non abordés en FLE.

Notons que les verbes de positionnement se répartissent dans plusieurs classes de notre liste (ancrage\_construction, communication, analyse<sup>3</sup> et processus cognitif). Nous avons choisi sept sous-classes représentatives. Notre analyse interroge la mise en relation entre l’expression de la prise de position de l’auteur et l’effacement énonciatif.

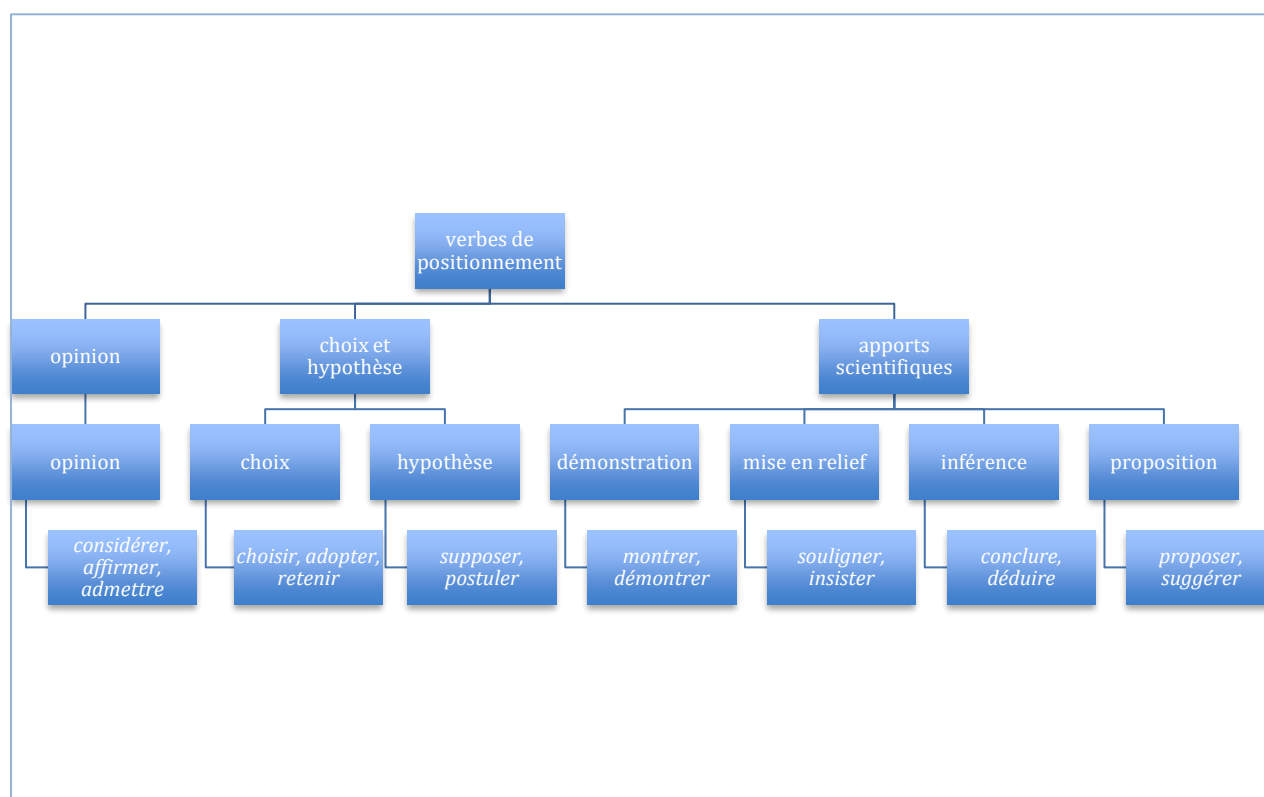
La figure 4-13 présente la classification des verbes de positionnement à étudier qui s’inspire de celle adoptée par Tutin (2010a). Dans son étude, elle a effectué une désambiguïsation sémantique des verbes de positionnement associés à un pronom sujet (auteur du texte) et a distingué trois grandes classes :

1. les verbes qui renvoient à une opinion ou à un questionnement (*considérer, se demander...*) ;
2. les verbes qui indiquent un choix ou une intention (*choisir, souhaiter...*) ;

---

<sup>3</sup> Les verbes de démonstration et les verbes d’inférence font partie des verbes d’analyse, mais étant donnée leurs fonctions rhétoriques et discursives (voir la figure 4-29), nous avons choisi de les rapprocher des verbes de positionnement et de les analyser dans cette section.

3. les verbes qui indiquent un apport spécifique de l'auteur (*montrer, proposer...*).



**Figure 4-13 : Classification des verbes de positionnement à étudier**

Contrairement à Tutin qui a intégré tous les verbes présentant un engagement fort de l'auteur dans l'écrit scientifique, nous n'avons retenu que les verbes du LST. Ainsi, les verbes renvoyant à un questionnement (*se demander*) et à une intention (*vouloir, souhaiter*), n'appartenant pas aux verbes du LST, ont été exclus de notre étude. Par ailleurs, les verbes indiquant une convergence/divergence (*se distinguer de, rejoindre*) ne sont pas intégrés pour deux raisons principales. D'une part, ces verbes présentent peu de variété au niveau des constructions syntaxiques ; d'autre part, ces verbes étant polysémiques, nous n'avons pas relevé une fréquence assez importante dans l'acception du positionnement. Comme le souligne Tutin (*ibid.* p. 31), il apparaît que la convergence/divergence est souvent exprimée à l'aide des adverbiaux tels que *contrairement à, à la différence de, à l'instar de*.

Parmi les verbes de positionnement retenus, les verbes indiquant un apport scientifique sont plus nombreux. Nous procédons maintenant à l'analyse linguistique de ces verbes.



### 4.4.1. Verbes d'opinion

Nous entendons par verbes d'opinion les verbes psychologiques qui partagent le sens de base 'avoir tel sentiment, telle activité consciente' (d'après le LVF). Ils sont notamment abordés dans l'étude de la modalisation, à côté des adverbess *à mon avis*, *selon moi*, etc. (Tuchais, 2013). Dans l'écrit scientifique, ils servent à introduire le jugement asserté par l'auteur, impliquant par conséquent sa prise en charge d'un contenu de pensée (Yan, 2014). Ces verbes sont « plus souvent introduits à l'aide d'un pronom incluant la communauté de discours (*nous* et *on* exclusifs) et sont souvent fortement modalisés » (*on peut considérer que*) (Tutin, 2010b, p. 38). Nous avons relevé une dizaine de verbes d'opinion (voir l'Annexe 2), mais nous n'abordons ici que trois verbes *considérer* (2548), *affirmer* (729) et *admettre* (379). Au niveau sémantique, ces trois verbes sont non ambigus ; au niveau syntaxique, ils présentent des constructions variées. Notons par ailleurs que le sens du verbe *affirmer* est différent de deux autres verbes : ces derniers induisent qu'une réflexion peut avoir lieu, tandis que dans l'affirmation, l'opinion est tranchée. Cependant, en termes de modalisation, il semble intéressant de les analyser ensemble.

#### a. Constructions syntaxiques et valeurs sémantiques

La construction prototypique de cette sous-classe est la structure en complétive *que* qui appelle souvent le modalisateur *pouvoir* [57].

[57] Les mêmes monnaies peuvent venir tantôt en équivalents à des biens matériels, tantôt en substituts à des humains. **On peut même affirmer que** leur spécificité est plutôt qu'elles soient, ou puissent être, équivalents (dans l'échange matériel) et substituts (dans l'échange social). (anthropologie.xml-s11800)

Notons que cette association est fréquente pour les trois verbes étudiés : *pouvoir considérer que* (359/2548), *pouvoir affirmer que* (51/729), *pouvoir admettre que* (34/379). Cette tendance à la modalisation peut également s'observer dans le contexte impersonnel *il est possible de* [58] :

[58] **Il est donc possible de considérer**, avec Jenkins et Perrow (1977), **que** la situation de « groupe déshérité » se traduit par un handicap du point de vue de la capacité à recourir

à des moyens d'action traditionnels tels que le vote, l'argent permettant de financer des campagnes ... (sociologie.xml-s2566)

Rappelons que nous avons abordé les marques d'atténuation lors de l'analyse des verbes d'interprétation. Ici, l'emploi des marques linguistiques (*on peut considérer, il est possible de*) a pour fonction principale d'atténuer l'assertion de l'auteur. Au lieu d'exprimer une opinion tranchée du type *je pense que*, l'auteur introduit son point de vue en favorisant la crédibilité de son discours, ce qui permet de persuader le lecteur de son bien-fondé (Hyland, 2005b ; Grossmann & Tutin, 2010b ; Rinck, 2010a).

Ensuite, nous avons remarqué que ces trois verbes apparaissent souvent dans des constructions causatives (*amener à considérer que, permettre d'affirmer/considérer que, conduire à admettre que*). Dans l'exemple [59], l'auteur justifie le bien-fondé de son opinion par les entretiens et l'analyse réalisés. De ce fait, son propos gagne davantage en crédibilité, impliquant un raisonnement causal et une dimension argumentative.

[59] Au-delà de ces différences, une série d'entretiens menés depuis 2004 avec des acteurs institutionnels, des urbanistes et des architectes, des militants associatifs et des acteurs locaux, ainsi que l'analyse qualitative d'un large corpus de documents [3] **nous permettent d'affirmer que** non seulement les pouvoirs publics ont fait de l'espace public un outil pour améliorer l'environnement des espaces centraux [4]... (geo.xml-s4092)

Ci-dessous en [60], la présence du gérondif (*en considérant (88/2548)/affirmant (37/729) que*) peut servir à introduire un point de vue lié à une activité scientifique, et constitue souvent une justification du choix :

[60] Nous définirons la ségrégation comme la « traduction de différences sociales dans l'espace », **en considérant qu'**elle « se manifeste dès que des individus, classés par la société dans des catégories sociales distinctes, dotées d'une valorisation sociale différenciée, se trouvent séparés dans l'espace et sont ainsi amenés à peu se côtoyer » (Delvaux, 2005, p. 276). (scienceseducation.xml-s99)

Notons que le verbe *admettre* se distingue des autres verbes par le fait qu'il est fréquemment utilisé pour introduire un point de vue reconnu dans la communauté

scientifique, souvent remis en cause par l'auteur. La fonction est de permettre à l'auteur de se démarquer des autres chercheurs en marquant ainsi l'apport de sa recherche : *on admet généralement que* (5), *Il est communément* (12)/*généralement* (5) *admis que* [61] :

[61] **Il est admis généralement que** la conception et la réalisation de PISA ont réduit ces biais au maximum de ce que les techniques actuelles autorisent en la matière. **Mais** les biais culturels qui demeurent sont-ils de nature à fausser l'appréhension des écarts de compétence entre pays, tels que ceux que nous mobilisons dans cet article? La réponse nous paraît négative. (scienceseducation.xml-s6523)

En ce qui concerne les adverbes, nous avons remarqué que ces verbes appellent souvent des adverbes à fonction discursive (Tran, 2014) tels que *par ailleurs*, *en outre*, *cependant*, *en effet*, introduisant un changement de perspective ou non sur le thème développé [62].

[62] **On considère, en effet, que** les producteurs sont caractérisés par un paramètre de productivité lié à la terre qu'ils possèdent. (eco.xml-s17449)

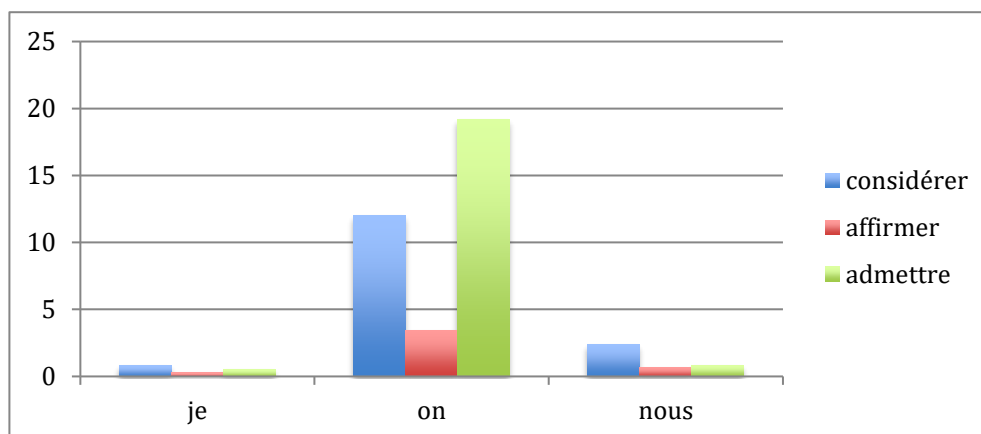
De manière générale, les constructions verbales de cette sous-classe se caractérisent par un positionnement fort et la tendance à la modalisation, le tableau 4.8 récapitule :

Fonctions	Constructions syntaxiques récurrentes
Formulation atténuée d'une opinion	<i>pouvoir considérer que</i> (359/2548), <i>pouvoir affirmer que</i> (51/729), <i>pouvoir admettre que</i> (34/379)
Justifier le bien-fondé de son opinion	Construction causative : <i>amener à considérer que</i> (8/2548), <i>permettre d'affirmer que</i> (9/729), <i>conduire à admettre que</i> (4/379)
Introduire un point de vue relié à l'activité	gérondif : <i>si l'on admet</i> (45/379)/ <i>considère</i> (23/2548) <i>que</i> , <i>en considérant</i> (88/2548) / <i>affirmant</i> (37/729) <i>que</i>
Introduire un fait reconnu (mais peut-être à remettre en cause)	<i>on admet généralement</i> (5/379) <i>que</i> , <i>il est communément</i> (12/379)/ <i>généralement admis que</i> (5/379)
Introduire un changement de perspective	co-occurrences avec les adverbes discursifs comme <i>par ailleurs</i> , <i>en outre</i> , <i>cependant</i>

**Tableau 4.8 : Récapitulatif des constructions syntaxiques récurrentes des verbes d'opinion et leurs fonctions correspondantes**

## b. Pronoms sujets mettant en jeu l'auteur

Les verbes d'opinion acceptent un sujet humain. Il nous semble donc important d'observer les associations entre les verbes d'opinion et les pronoms sujets mettant en jeu l'auteur (*je*, *on* et *nous*) afin d'étudier la dimension énonciative liée à cette sous-classe.



**Figure 4-14 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes d'opinion)**

Comme le montre la figure 4-14, il apparaît que l'emploi des pronoms *nous* et *je* reste marginal. Par ailleurs, il est intéressant de noter que le verbe *affirmer* est peu utilisé avec ces trois pronoms. En effet, ce verbe est fréquemment associé au sujet renvoyant à un autre chercheur [63] :

[63] **Dervin** pose d'autres éléments qui rejoignent les prémisses conceptuelles de Weick. **Elle affirme qu'**une des découvertes du sense – making est que l'acte de communication et de partage des connaissances fonctionne mieux quand les individus s'ancrent dans leurs histoires et leurs cadres conceptuels (**Dervin, 1998** ; Savolainen, 2006). (scinfo.xml-s2383)

Cet emploi particulier peut être expliqué par le sens de ce verbe qui se rapporte à une opinion tranchée et est davantage utilisé pour rapporter les discours d'autrui.

Pour les verbes *considérer* et *admettre*, le pronom *on* est préféré, en particulier dans l'association *on+pouvoir*, ce qui traduit une responsabilité atténuée de l'auteur.

[64] Il s'agit donc bien d'une relation intersubjective où chacun des protagonistes est

reconnu en tant que sujet désirant et **l'on peut affirmer que** cette dimension est effectivement présente dans la relation Nathanaël/Alec... (scinfo.xml-s9240)

Comme pour le verbe *interpréter*, le *on* semble mettre en place une négociation. La fonction principale est de faire partager son point de vue par la communauté scientifique et de laisser au lecteur « la possibilité d'y adhérer ou non » (Hyland 2005b, cité par Grossmann & Tutin, 2010b, p. 3).

L'analyse des trois verbes d'opinion permet de constater que l'expression de l'opinion est fortement modalisée, à travers notamment *on+pouvoir*. Par ailleurs, étant donné que l'auteur souhaite faire accepter ses idées, les verbes d'opinion mettent en jeu des relations de preuve, où l'opinion est introduite avec une justification fondée scientifiquement (ex : construction causative). Enfin, ces verbes se caractérisent par leur co-occurrence avec les adverbes à fonction discursive (ex : *cependant*, *en outre*, etc.), ce qui témoigne d'une dimension argumentative fortement présente. Les patrons lexico-syntaxiques de cette sous-classe sont présentés dans le tableau 4.9 :

considérer	1. [[humain=auteur, chercheur]] (pouvoir) <i>considère</i> (adverbe discursif) [complétive <i>que</i> ] Ex : <b>On peut considérer que</b> la réponse construite par P2 est pire que celle construite par P1 parce que c'est elle qui porte le biais... (ling.xml-s10017) Sens : Quelqu'un se forme une opinion autour d'un fait scientifique ou quelque chose d'abstrait.
affirmer	1. [[humain=auteur, chercheur]] (pouvoir) <i>affirme</i> (adverbe discursif) [complétive <i>que</i> ] Ex : <b>On peut donc affirmer que</b> les individus caractérisés par ces trois profils sociaux ont une probabilité très faible d'avoir des amis virtuels. (scinfo.xml-s4512) 2. [[relation cas processus]] permet d'affirmer [complétive <i>que</i> ] Ex : Les <b>résultats permettent d'affirmer qu'</b> un système de cinq modalités axiologiques préside au jugement de ressemblance... (psycho.xml-s12293) 3. comme l'affirme [[humain=chercheur]] Ex : <i>D'autre part, son élaboration n'a rien d'arbitraire, comme l'affirme Bruno Latour.</i> (sociologie.xml-s11752) Lexset [[relation cas processus]] : <i>résultat, analyse, fait, condition</i> Sens : Quelqu'un présente un fait scientifique comme vrai.
Admettre	1. [[humain=chercheur]] (pouvoir) <i>admet</i> (adverbe discursif) [complétive <i>que</i> ] Ex : <b>On peut admettre à l'inverse, avec H. Renchon (1967) qu'</b> elle constitue une prise à témoins de la "galerie"... (ling.xml-s1085) Sens : Quelqu'un accepte l'idée que quelque chose est vrai ou possible.
Valeurs sémantiques	valeur argumentative, valeur interprétative, valeur évaluative
Routines	<i>Il est/semble possible d'affirmer que</i> (6/729) <i>il faudrait/faut admettre que</i> (17/379) <i>si l'on admet</i> (45/379)/ <i>considère</i> (23/2548) <i>que</i>

	<i>en considérant (88/2548)/affirmant (37/729) que Il est communément (12/379)/généralement (5/379) admis que amener à considérer que (8/2548) conduire à admettre que (4/379)</i>
Commentaires	Le sujet est souvent le pronom <i>on</i> , occurrences avec des adverbes discursifs tels que <i>par ailleurs, en outre, cependant, en effet</i>

**Tableau 4.9 : Patrons lexico-syntaxiques des verbes d'opinion**

#### 4.4.2. Verbes de démonstration

Dans le LVF, les verbes de démonstration sont classés comme des « verbes de communication » et ont le sens de base ‘faire connaître la preuve’. Ces verbes consistent à ‘établir au moyen de faits, de témoignages, de raisonnements, la vérité ou la réalité de (quelque chose)’ (d’après le TLFi). Deux verbes représentatifs de cette sous-classe se distinguent nettement en nombre d’occurrences : *montrer* (3114) *démontrer* (326). Dans l’écrit scientifique, la vérité se construit à travers des données scientifiques ou d’autres propositions, justifiées comme vraies. Ainsi, la mise en scène de preuve joue un rôle important pour établir l’assertion de l’auteur dans la recherche scientifique.

##### a. Constructions syntaxiques et valeurs sémantiques

Nous relevons trois constructions syntaxiques principales : la construction avec complétive (*on montre que/comment*), la construction avec complément nominal (*la recherche démontre l’importance*) et la construction en *comme* (*comme on l’a montré*). Cette dernière est moins fréquente que les deux premières alors que la construction avec complétive est la plus productive. Ces constructions comportent un sujet humain et prennent le sens de ‘justifier certains faits scientifiques avec des arguments’.

La construction avec complétive met en scène le fait prouvé soit par d’autres chercheurs [65], soit par l’auteur lui-même [66] (prouvée ou à prouver). Quand il s’agit de l’auteur, la construction comporte souvent un déictique de lieu (*ici, plus haut*), le lecteur étant invité à s’y reporter. Cette caractéristique présente des points communs avec les verbes de constat et d’autres verbes d’analyse ayant une valeur métatextuelle comme *étudier, examiner, décrire*. Dans l’exemple [66], le verbe introduit une démonstration servant à construire l’argumentation de l’auteur.

[65] D'autres études expérimentales avaient déjà nuancé cette théorie. **Martin (1979) a montré qu'**en fonction de la densité des éléments locaux d'une forme hiérarchique, l'effet de dominance du global sur le local peut s'inverser. (psycho.xml-s10204)

[66] **Je tenterai ici de démontrer que** l'idée développée par Pirenne doit être révisée. (histoire.xml-s11496)

La valeur métatextuelle (en particulier pour le verbe *montrer*) est mise en évidence lorsque le verbe s'associe à un adverbe d'énumération (ex. : *d'abord, dans un premier temps*), notamment dans l'introduction pour annoncer le plan [67] ou dans la conclusion pour résumer les points importants abordés [68]. Ici, le verbe *montrer* se rapproche du verbe *présenter*.

[67] **Dans un second temps, nous montrons que** le choix et la combinaison des instruments privilégiés contribuent parfois à l'extension du champ d'intervention de l'UE...(sciences politiques\_148)

[68] Cependant, **nous avons montré qu'**il existe une zone où l'introduction de la concurrence conduit à la présence de trop de services offerts par rapport à une situation de premier rang. (économie\_232)

On rencontre également un nombre non négligeable de compléments nominaux qui est généralement abstrait, en particulier pour le verbe *montrer* qui est très productif. Ces compléments nominaux seront décrits plus précisément plus bas dans la partie sur les arguments nominaux.

Par ailleurs, il faut souligner la présence importante de la construction en *comme* qui comporte un verbe au passé composé, au passif réduit, ou un verbe au présent : *comme nous l'avons déjà montré, comme le démontre l'auteur*. Comme nous l'avons souligné dans le chapitre III, cette construction permet d'une part, à l'auteur de faire référence aux arguments précédents [69] ; et d'autre part, de prouver la véracité de ses propos, mettant en avant une forme d'autopromotion [70] :

[69] On serait, donc, tentée d'en déduire que les données contextuelles d'ordre situationnel n'ont d'incidence que discursive reposant sur une disjonction provisoire entre signe et situation référentielle, et de conclure que, **comme le démontre C. Kerbrat -**

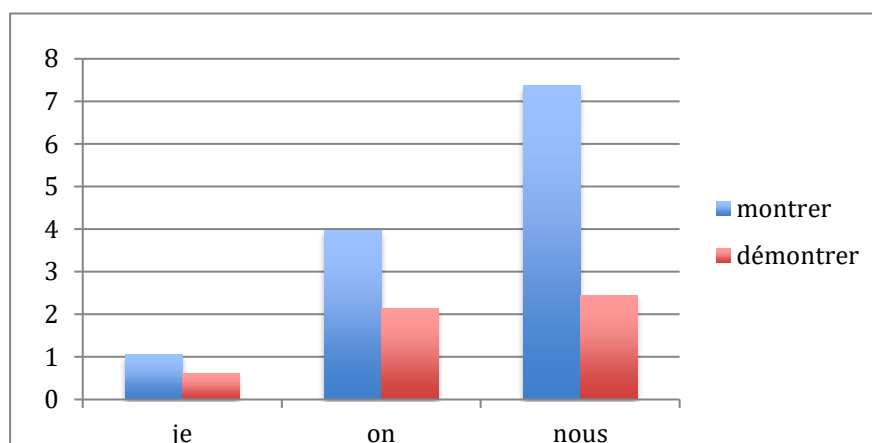
Orecchioni (1980) à propos des déictiques : "ce qui varie avec la situation, c'est le référent et non pas le sens". (ling.xml-s4061)

[70] Plusieurs différences persistent selon les zones prises en compte. **Comme nous l'avons démontré précédemment** [39], certaines régions ont vu le taux d'érosion de leur propre langue régionale (basque, corse, kanak, occitan, alsacien, breton, créoles) se ralentir, alors que d'autres continuent à assister à une chute exponentielle. (sciencespolitiques.xml-s8322)

D'autres constructions nous paraissent importantes du fait de leur fréquence dans le corpus d'analyse : *il a été montré que* (8/3114 occurrences), *permettre de montrer que* (80/3114), *permettre de démontrer* (8/326 occurrences), *tendre à démontrer que* (7/326), *(cela est) démontré* (24/326). Ces informations linguistiques doivent être abordées pour aider les apprenants à s'initier à la rédaction des écrits universitaires.

#### b. Pronoms sujets mettant en jeu l'auteur

Concernant les pronoms sujets, la figure 4-15 donne un aperçu de leurs fréquences relatives dans le corpus :



**Figure 4-15 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de démonstration)**

De manière générale, les verbes de démonstration sont plus fréquemment associés au pronom *nous*. On peut constater que le verbe *montrer* a une préférence marquée pour le pronom *nous*, renvoyant à l'auteur (singulier ou collectif) [71] :

[71] Du point de vue du taux de participation des usagers et du nombre de corrections



effectuées dans les bases de données, **nous avons montré que** les projets collaboratifs mis en Suvre par BAC sont faiblement efficaces. (scinfo.xml-s10938)

Dans l'exemple [71], la co-occurrence *nous + montrons* se rapporte souvent au processus d'écriture, avec par conséquent une valeur métatextuelle.

Lorsque les verbes sont associés au pronom *on*, ce dernier semble représenter seulement l'auteur (singulier ou collectif). Contrairement aux verbes d'opinion qui peut être employés avec le *on* inclusif, les verbes de démonstration n'impliquent pas une négociation avec le lecteur [72].

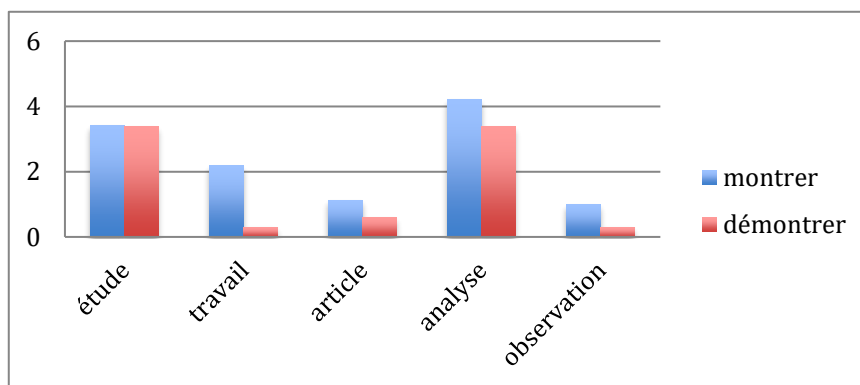
[72] Aussi s'intéresse-t-on spécifiquement à la pénalité monétaire devant être mise en place lorsque des pénalités budgétaires sont impossibles à pratiquer. **On montre** alors, sous des hypothèses admissibles pour les paramètres, **qu'**une Union monétaire où la politique budgétaire est décentralisée doit être plus inflationniste qu'une Union où la politique budgétaire est centralisée. (eco.xml-s15324)

Notons par ailleurs que le cas est différent pour le verbe *démontrer* qui marque relativement une faible proportion dans les associations « pronoms sujets+verbe ». En revanche, il s'intègre souvent dans des constructions au passif [73] :

[73] À ce moment, l'exactitude des hypothèses sur la validité des estimations **est démontrée** en produisant ces mêmes estimations de validité comme un produit de la solution. (sciencespolitiques\_46)

### c. Arguments nominaux

Au niveau des arguments nominaux, on remarque que les verbes de démonstration peuvent accepter des sujets métonymiques tels que *étude, travail, article, analyse, observation*, etc. L'auteur s'efface derrière l'ouvrage qu'il construit ou le processus cognitif impliqué (*cette analyse montre = nous montrons dans l'analyse*) (Tutin, 2010b). Étant donné le nombre non négligeable des sujets métonymiques associés à ces deux verbes, nous calculons leur proportion illustrée ci-dessous :

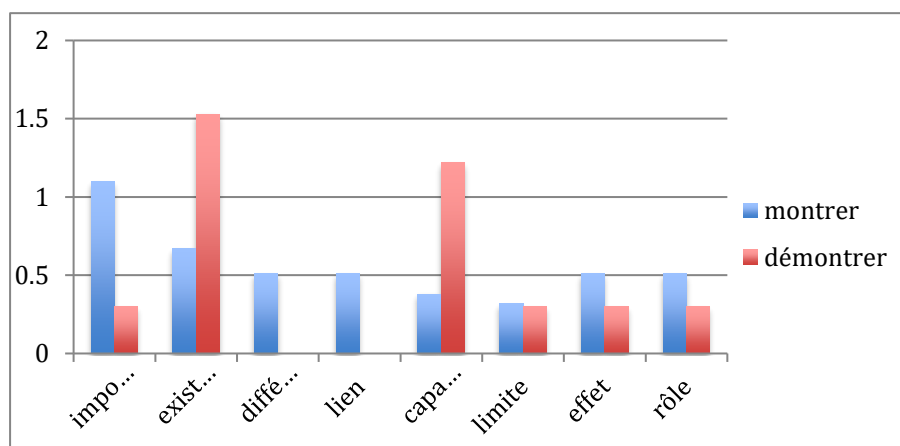


**Figure 4-16 : Les sujets métonymiques des verbes de démonstration (en %)**

Le nombre relevé est assez limité, mais il est intéressant de noter que le verbe *montrer* est celui qui s'associe le plus souvent à un nom métonymique, qui représente souvent l'étude d'un autre chercheur. En se fondant sur les résultats de l'étude d'autrui, l'auteur renforce son argumentation [74] :

[74] **Comme l'ont montré notamment les analyses** de Donzelot (2007), on assiste à un renouveau des interrogations sur le lien social : dans un contexte de forte croissance des inégalités sociales et territoriales, de développement de politiques centrées sur l'autonomie des acteurs. (scienceseducation.xml-s12298)

Les arguments nominaux de l'objet, contiennent des noms abstraits tels que *importance, existence, différence, lien, capacité, limite, effet, rôle* qui peuvent être regroupés dans les types sémantiques [[propriété]] et [[relation]]. La figure ci-dessous détaille, pour les deux verbes étudiés, les objets les plus fréquents :



**Figure 4-17 : Les noms objets des verbes de démonstration (en %)**

Dans la figure 4-17, les objets nominaux sont plus variés pour le verbe *montrer*

alors que le verbe *démontrer* a une préférence marquée pour les noms *existence* et *capacité*. Ceci pourrait s'expliquer d'une part par la différence de leurs fréquences dans le corpus et d'autre part par leurs différences sémantiques. En effet, comme déjà souligné, le verbe *montrer* peut revêtir une valeur métatextuelle où la dimension argumentative est absente. Cependant, cette valeur métatextuelle est presque absente pour le verbe *démontrer* qui implique toujours un processus d'explication et de raisonnement rigoureux. Comparons les exemples [75] et [76] :

[75] Après avoir montré la manière dont s'établissent les dynamiques coalitionnelles et les priorités [...] Puis je **montrerais l'importance** des mécanismes de concertation institués pour favoriser l'obtention de compromis. (sciencespolitiques.xml-s13663)

[76] Cela nous amène donc à privilégier une interprétation des présents résultats en termes d'efficacité intellectuelle. En effet, invoquer le processus de différenciation des aptitudes avec l'âge à titre de piste causale imposerait d'en **démontrer l'existence** effective au cours de l'enfance et, dans le même temps, d'expliquer pourquoi la relation négative entre QI et facteur général apparaît chez l'adulte non retardé. (psycho.xml-s9385)

Il nous semble donc que ces deux verbes synonymes utilisés dans le sens de 'démonstration scientifique' partagent des constructions similaires. Cependant, le verbe *montrer* a fondamentalement un rôle moins argumentatif étant donné sa valeur métatextuelle. Les deux verbes ont une tendance de modalisation avec le modalisateur *pouvoir*, mais la forme modalisée *on+pouvoir* reste moins fréquente par rapport aux verbes d'opinion. En effet, les verbes de démonstration s'associent préférentiellement aux pronoms qui renvoient à l'auteur (singulier ou collectif). Le récapitulatif des patrons est présenté dans le tableau ci-dessous:

montrer	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] (pouvoir) montre (adverbe discursif temps espace) [complétive <i>que</i> ou <i>comment</i>] Ex : <b>Martin (1979) a montré qu'en fonction de la densité des éléments locaux d'une forme hiérarchique, l'effet de dominance du global sur le local peut s'inverser.</b> (psycho.xml-s10204)</li> <li>2. [[humain=auteur, chercheur]] (pouvoir) montre (adverbe discursif temps espace) [[qualité relation]] Ex : <b>Nous espérons également avoir montré l'importance du positionnement séquentiel des questions réponses.</b> (ling.xml-s10319)</li> <li>3. comme le montre [[humain=auteur, chercheur]] (adverbe temps espace) Ex : <b>Comme on l'a montré plus avant, les déclarations de C. Vanneste entraînent la constitution de pôles antagonistes au sein même de l'UMP.</b></li> </ol>
---------	---

	<p>(sciencespolitiques.xml-s10059)</p> <p>4. [[entité abstraite événement]] permet de montrer [complétive que ou comment] Ex : <i>Cette <b>discontinuité</b> permet de montrer que l'équilibre est génériquement déterminé dès lors que cette proportion est strictement positive.</i> (eco.xml-s14052)</p> <p>5. [[entité abstraite événement]] permet de montrer [[qualité relation]] Ex : <i><b>L'analyse</b> détaillée des résultats obtenus nous a <b>permis de montrer l'existence</b> de cet effet et de le caractériser.</i> (psycho.xml-s11888)</p> <p>Lexset [[qualité relation]] : <i>importance, effet, différence, lien, existence, capacité, limite, rôle</i> Sens : Quelqu'un justifie et met en évidence la vérité (ou l'existence) d'un fait scientifique.</p>
démontrer	<p>1. [[humain=auteur, chercheur]] (pouvoir) démontre (adverbe discursif) [complétive que ou comment] Ex : <i><b>ROCHET et CHONÉ</b> [1998] <b>démontrent</b> en effet <b>que</b> le monopole multiproduit utilise généralement les ventes liées.</i> (eco.xml-s11192)</p> <p>2. [[humain=auteur, chercheur]] (pouvoir) démontre (adverbe discursif) [[qualité relation]] Ex : <i>Enfin, <b>nous</b> avons également <b>démontré les limites</b> d'une approche fondée sur des critères de surface...</i> (ling.xml-s13545)</p> <p>3. comme le démontre [[humain=auteur, chercheur]] Ex : <i>Notons que ce résultat, obtenu dans le cadre de notre modèle avec demande unitaire, est plus général, <b>comme l'a démontré JULLIEN</b> [2000].</i> (eco.xml-s11308)</p> <p>4. [[entité abstraite événement]] permet de démontrer [[qualité relation]] Ex : <i><b>L'étude</b> de cette liste <b>permet de démontrer l'hétérogénéité</b> de la ligue qui comprend à la fois des représentants de l'élite marchande et immobilière...</i> (histoire.xml-s24468)</p> <p>Lexset [[qualité relation]] : <i>importance, effet, existence, capacité, limite, rôle</i> Sens : Quelqu'un fournit la preuve pour établir la vérité (ou l'existence) d'un fait scientifique en suivant un raisonnement rigoureux.</p>
Valeurs sémantiques	valeur argumentative, valeur démonstrative, valeur métatextuelle
Routines	<p><i>il a été montré que</i> (8/3114)  <i>tendre à démontrer que</i> (7/326)  <i>(cela est) démontré</i> (24/326)  <i>l'objectif est de montrer</i> (19/3114)</p>
Commentaires	Sujets métonymiques sont fréquents, le passif est présent, en particulier pour le verbe <i>démontrer</i> , le pronom <i>nous</i> est plus fréquent pour le verbe <i>montrer</i> . Le verbe <i>montrer</i> s'associe fréquemment aux adverbes d'énumération ( <i>d'abord, dans un premier temps</i> ) de temps ( <i>précédemment</i> ) ou d'espace ( <i>ici</i> ).

Tableau 4.10 : Patrons lexico-syntaxiques des verbes de démonstration

#### 4.4.3. Verbes de mise en relief

Dans l'écrit scientifique, les verbes de mise en relief remplissent la fonction d'attirer l'attention du lecteur sur « un aspect saillant ou imprévu du sujet dont il est question » (Siepman 2007, p. 129). La mise en relief s'inscrit dans le champ sémantique de positionnement, établissant la relation entre l'auteur (rôle du chercheur) et l'élément thématique auquel l'auteur accorde une importance particulière. Nous avons

relevé deux verbes de mise en relief tels que *souligner* (1158) et *insister* (408) dans le corpus. Leur sens repéré dans le LVF est de ‘communiquer quelque chose avec force’. En plus de ces deux verbes, nous pouvons également citer des expressions polylexicales ayant la même valeur sémantique : *mettre en lumière*, *mettre en évidence* et *mettre en exergue*.

### a. Constructions syntaxiques et valeurs sémantiques

Le verbe transitif direct *souligner* s’associe essentiellement à deux constructions : la première comporte un complément nominal [77] alors que la deuxième est celle de la complétive *que* [78] :

[77] **Knobloch souligne l’importance** du langage intérieur pour la production des énoncés langagiers et pour leur compréhension...(Knobloch 1984 : 230). (ling.xml-s7301).

[78] Bien au contraire, ces incompétences relèvent d’un désintérêt pour ces questions, bien compréhensible une fois le contexte de la plupart de leurs pratiques rappelé. En effet, **nous avons souligné** à plusieurs reprises **que** celles-ci concernent essentiellement la sociabilité. (scinfo.xml-s10770)

Quant au verbe transitif indirect *insister* suivi par la préposition *sur*, il peut prendre un complément nominal ou apparaître dans la construction comportant *le fait que* [79].

[79] **Nous avons** pour notre part bien **insisté sur le fait que** les instruments n’apparaissent pas dans un vide institutionnel et nous avons donc privilégié une définition englobante de l’instrumentation... (sciencespolitiques.xml-s12108)

Ces constructions constituent un moyen prototypique de signaler un fait important. On peut les rapprocher des constructions verbales associées au verbe *noter* (*on peut noter que*, *il est à noter que*, *il faut noter que*) qui consistent à attirer l’attention du lecteur sur un fait saillant.

La présence du gérondif peut introduire une précision et une manière. Dans l’exemple [80], *en insistant* apporte une précision sur la façon dont on va mener l’analyse :

[80] Ces deux secteurs seront analysés successivement, **en insistant** pour chacun d’entre eux **sur** le choix et la combinaison des instruments, d’une part, et sur leurs effets pour le développement de l’action publique européenne, d’autre part. (sciencespolitiques.xml-s17219)

Par ailleurs, il est fréquent de voir que ces deux verbes mobilisent un adverbe discursif : *par exemple, ainsi, également, d’ailleurs, aussi*, avec par conséquent une dimension argumentative bien présente [81] :

[81] Son livre, *A Question of Torture* (2006), montre que la « brutalisation » de la torture s’est accrue grâce aux travaux approfondis de psychologues (et non pas de psychiatres) américains et canadiens dans les décennies 1950 – 1960... [26] **On soulignera cependant que** cette approche visant à « gagner les cSurs et les esprits » s’inspire du programme « Cords », appliqué par les États-Unis au Sud - Viêt-Nam dans les années 1970. [27]. (anthropologie.xml-s4763)

Dans cet exemple, l’auteur se situe par rapport à l’étude d’un autre chercheur, mettant en relief un changement de point de vue. Ici, la mise en relief est renforcée par le futur injonctif, caractéristique similaire au verbe *noter* soulignée par Grossmann et Tutin (2010b).

En dehors de ces constructions en commun, le verbe *souligner* peut également accepter la construction au passif et la construction en *comme* (141/1158 occurrences). Il est intéressant de constater que ce verbe est fréquemment utilisé pour signaler une citation ou renvoyer le lecteur à d’autres études qui servent de l’argument de l’auteur [82]. Précisons que cette fonction est associée aux verbes *expliquer, montrer, affirmer* analysés précédemment. Par ailleurs, il se distingue du verbe *insister* par une valeur métatextuelle, comme l’illustre l’exemple [83] :

[82] En effet, **comme l’ont souligné** notamment CAMPBELL et SHILLER [1991], la pente observée et la pente théorique (calculée en tenant compte des contraintes imposées par la théorie des anticipations) présentent souvent des évolutions très proches... (eco.xml-s10084)

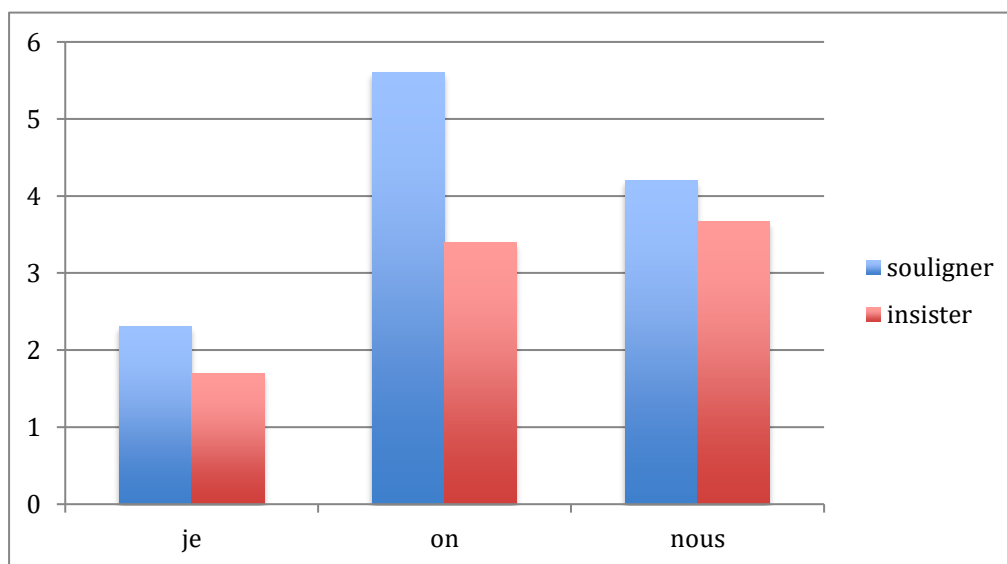
[83] **Comme nous l’avons souligné** dans la première partie de cette étude, bien que certaines recherches restituent des relations empiriques significatives, l’absence de

relation entre l'engagement affectif et les comportements de citoyenneté organisationnelle est une réalité empirique... (psycho.xml-s1856)

Dans cet exemple, l'auteur fait un rappel d'un point important en assurant une cohérence textuelle.

### b. Pronoms sujets mettant en jeu l'auteur

Comme on peut le voir dans le graphique 4-18, les pronoms mettant en jeu l'auteur ne sont pas fréquents. Ceci pourrait s'expliquer par le fait que ces deux verbes sont fréquemment utilisés pour citer des travaux autres que ceux de l'auteur.



**Figure 4-18 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de mise en relief)**

Les pronoms *on* et *nous* sont utilisés à peu près dans la même proportion. Dans les deux cas, les pronoms représentent notamment l'auteur (singulier ou collectif) parce que l'acte de mise en relief implique un patient (le lecteur) implicite à qui l'auteur souhaite signaler quelque chose d'important. Il semble qu'ici l'emploi inclusif du pronom *on* est absent, voir l'exemple [84].

[84] Loin de souscrire à une forme de déterminisme institutionnel, **nous préférons insister**, à la suite de Pierson, **sur** les limites d'une analyse par trop fonctionnaliste qui ferait l'impasse sur l'horizon temporel réduit ou encore sur les effets indésirables qui caractérisent les moments de création institutionnelle [70].

(sciencespolitiques.xml-s16012)

En dehors des pronoms mettant en jeu l'auteur, on remarque aussi l'emploi métonymique au niveau du sujet que nous allons présenter ci-dessous.

### c. Arguments nominaux

Nous avons constaté que les noms métonymiques les plus fréquemment mobilisés par les deux verbes sont *étude*, *rapport*, *travail* et *recherche*. Précisons que des sujets métonymiques similaires ont été relevés pour le verbe *montrer*. Contrairement aux verbes d'opinion qui « mettent en jeu un processus cognitif complexe » et qui acceptent difficilement les sujets métonymiques (Tutin 2010, p. 36), les verbes de mise en relief, ayant une valeur métatextuelle, semblent avoir plus d'affinité avec les sujets métonymiques (*cette étude souligne l'importance de*). Par ailleurs, il est intéressant de noter qu'ici, l'emploi métonymique est utilisé notamment pour renvoyer aux travaux d'autrui. En se référant à d'autres études, l'auteur exprime soit une convergence [85] soit une divergence [86] sur laquelle se construit le point de vue de l'auteur.

[85] On remarque que la navigation est nettement plus significative : elle est sept fois supérieure à la recherche directe. Les usagers consultent donc les articles essentiellement dans le cadre d'une navigation. **Une étude récente**, qui a également **souligné cet aspect** [17], présente clairement dans son analyse l'importance de l'activité de navigation chez les usagers des plates-formes de revues électroniques. (scinfo.xml-s13101)

[86] Pourtant, dans la littérature consacrée aux discours sur l'homme appuyés sur les sciences de la vie au XIXe siècle, cet aspect du succès de la biologie est nettement moins souligné que d'autres. **Ces travaux insistent en effet généralement davantage sur** le registre essentialiste, naturaliste et causal de ces discours. Ainsi a-t-on particulièrement prêté attention à tout ce qui souligne dans ces théories les supposées causes « organique » des différences entre humains... (sciencespolitiques.xml-s11)

Au niveau de l'objet, il nous semble que les deux verbes ont tendance à choisir des noms renvoyant à la qualité d'une entité abstraite ou concrète (*importance*, *rôle*,



*aspect, intérêt, difficulté*). Il s'agit donc d'un aspect particulier que la personne souhaite mettre en relief.

En outre, nous remarquons une fréquence très élevée du nom *fait* associé au verbe *insister*, ce qui est lié à la présence de la construction *insister sur le fait que*. La répartition des noms est illustrée dans la figure 4-19 :

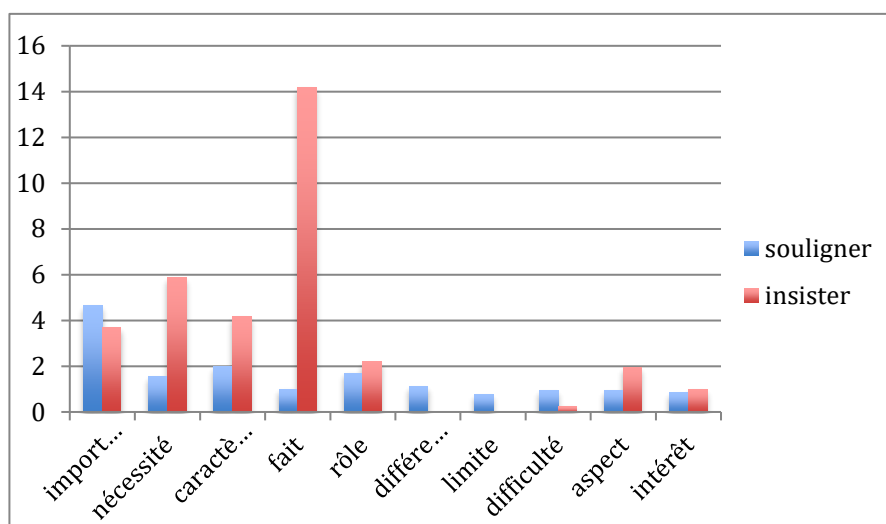


Figure 4-19 : Les noms objets des verbes de mise en relief (en %)

Enfin, il nous semble que cette sous-classe partage des propriétés sémantique et syntaxique similaires au verbe *montrer* (constructions syntaxiques, emploi métonymique, l'emploi des pronoms *on* et *nous*). Nous avons montré que les verbes de mise en relief remplissent deux fonctions essentielles dans l'écrit scientifique : mettre en relief un aspect saillant lié à un fait scientifique et signaler une citation afin de renforcer l'argument de l'auteur. Nous présentons les patrons prototypiques des verbes *souligner* et *insister* dans le tableau ci-dessous :

souligner	<ol style="list-style-type: none"> <li>[[humain=auteur, chercheur]] (pouvoir) souligne (adverbe discursif) (le fait) [complétive <i>que</i>] Ex : <b>L'auteur souligne que</b> l'idée de « compétence collective », dans la littérature, est peu présente...(scinfo.xml-s11064)</li> <li>[[humain=auteur, chercheur]] souligne (adverbe discursif) [[qualité]] Ex : Enfin, <b>on soulignait la nécessité</b> d'élargir le contenu du cours sur l'histoire du Québec...(scienceseducation.xml-s3410)</li> <li>comme le souligne [[humain=auteur, chercheur]] Ex : <b>Comme nous l'avons souligné ci-dessus</b>, nous considérons pour simplifier que les crédits à la clientèle ne sont pas négociables sur un marché. (eco.xml-s14640)</li> </ol> <p>Lexset [[qualité]] : <i>importance, caractère, nécessité, rôle, aspect, intérêt</i> Sens : Quelqu'un attire l'attention du lecteur sur un fait ou une qualité liée à quelque chose avec insistance particulière.</p>
-----------	---

insister	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] insiste (adverbe discursif) sur le fait que Ex : <b>On insiste par exemple sur le fait que</b> la maternité corporelle n'est pas définie de façon exclusivement anténatale [22]. (anthropologie.xml-s5885)</li> <li>2. [[humain=auteur, chercheur]] insiste (adverbe discursif) sur [[qualité]] Ex : <b>Il insiste sur le rôle</b> que l'usine joue sur le marché local du travail et n'hésite pas à se définir comme "l'homme de la classe ouvrière". (sciencespolitiques.xml-s16241)</li> </ol> <p>Lexset [[qualité]] : importance, caractère, nécessité, rôle, aspect, intérêt, différence, limite, difficulté</p> <p>Sens : Quelqu'un accorde de l'importance à un fait ou une qualité liée à quelque chose parce qu'il va peut-être développer cette idée dans un examen détaillé.</p>
Valeurs sémantiques	valeur argumentative, valeur métatextuelle,
Routines	<i>être souligné</i> (35/1158) <i>il a été souligné que</i> (4/1158) <i>il faut souligner que</i> (45/1158) <i>il faut s'insister sur</i> (11/408)
Commentaires	sujets métonymiques présents, le gérondif est fréquent, co-occurrences avec les adverbes discursifs tels que <i>par exemple, ainsi, également, aussi,</i>

**Tableau 4.11 : Patrons lexico-syntaxiques des verbes de mise en relief**

#### 4.4.4. Verbes de choix

Les verbes de choix jouent un rôle important dans la présentation de la démarche méthodologique. Ces verbes à deux arguments (sujet humain et objet inanimé) contribuent à relier l'auteur (rôle du chercheur) aux « objets scientifiques aussi divers qu'une *démarche*, une *méthodologie* ou un *cadre théorique* issus des travaux des prédécesseurs ou des pairs » qui sont à intégrer dans la recherche (Grossmann *et al.*, 2009, p. 187). Nous avons retenu trois verbes représentatifs : *choisir* (837), *adopter* (664), *retenir* (640). Dans le LVF, ces verbes sont classés comme verbes de saisie ou de possession, ayant le sens de base 'prendre quelque chose d'abstrait'. La première observation du corpus nous révèle qu'ils apparaissent souvent dans le développement, en particulier la partie méthodologique pour marquer le choix ou l'emprunt de l'auteur.

##### a. Constructions syntaxiques et valeurs sémantiques

La structure standard comporte un complément nominal, mais nous avons constaté que ces verbes apparaissent plus fréquemment dans la construction au passif (en particulier le passif réduit) [87] et [88] :

[87] Si l'**approche adoptée** ici est néanmoins intéressante pour mettre en évidence des

différences entre pays, elle demandera à l'évidence à être prolongée par des approches plus qualitatives et/ou multi - niveaux, si cette exploration suggère qu'il y a là des perspectives stimulantes pour mieux comprendre les systèmes éducatifs et les transformations qu'ils induisent chez les élèves. (scienceseducation.xml-s12341)

- [88] Le modèle théorique de l'étude a été comparé avec les trois modèles alternatifs potentiels. **Le modèle final retenu** a ensuite été utilisé pour vérifier les hypothèses formulées dans la partie théorique. (psycho.xml-s3596)

Dans ces deux exemples, les objets scientifiques (*approche, modèle*) sont thématiques et deviennent le sujet de la phrase. L'auteur met l'accent sur l'intérêt de son choix [87] et l'usage de ces objets [88]. En fait, il est intéressant de noter que ces verbes apparaissent dans un co-texte avec des marques linguistiques de l'objectif ou de la cause, mettant en jeu une justification du choix. Par exemple, on rencontre l'adverbe, *donc*, les connecteurs *pourquoi, car* et les prépositions *pour, afin de*, comme l'illustre l'exemple [89] :

- [89] Le but du présent article est d'**adopter une démarche** similaire pour le modèle WS - PS développé par CAHUC et ZYLBERBERG [1999] (désormais référencé CZ) **afin de** mieux comprendre comment, dans le cadre de ce modèle, le système d'indemnisation du chômage affecte le taux de chômage de long terme. (eco.xml-s1039)

Le gérondif est aussi présent pour les trois verbes *choisir* (25/837), *adopter* (28/664) et *retenir* (11/640). Il introduit souvent une manière ou une cause et semble confirmer que les verbes de choix sont fortement liés au raisonnement causal, comme dans les exemples [90] et [91] :

- [90] **En choisissant cette méthode**, nous avons postulé que les données verbales obtenues sont traitées comme une production monologale. (psycho.xml-s11157)
- [91] **En adoptant une perspective** de longue durée, nous pouvons comprendre les raisons pour lesquelles les autorités françaises ont accepté en 2007 de confier la définition des pratiques Snologiques à la Commission européenne, alors qu'elles s'y étaient fermement opposées en 1998. (sciencespolitiques.xml-s18437)

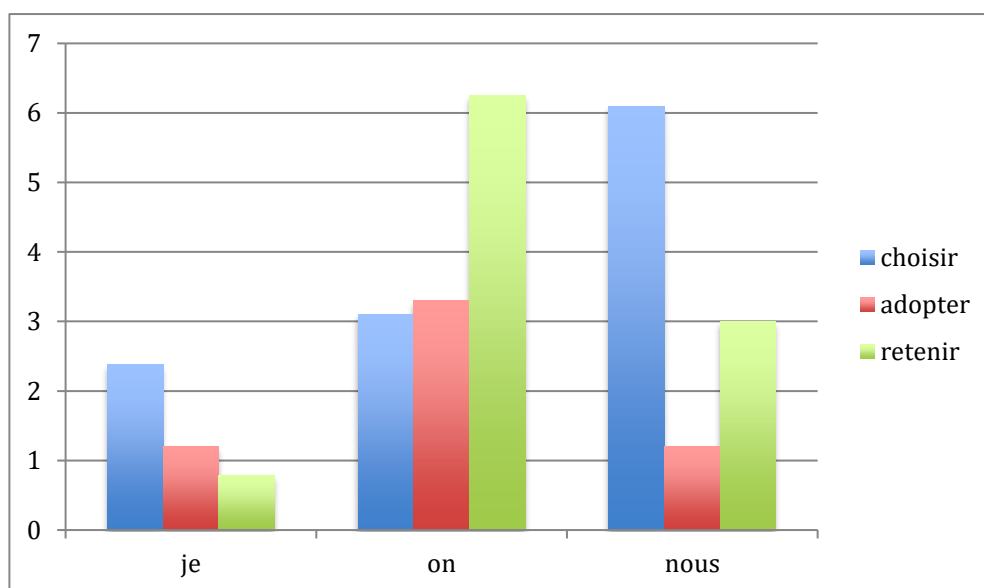
En dehors de ces constructions, le verbe *choisir* est le seul à privilégier une construction infinitive introduite par la préposition *de*. Les verbes infinitifs les plus fréquemment associés sont souvent des verbes d'analyse (*examiner, traiter, utiliser, présenter, etc.*). De ce fait, cette construction est utilisée de préférence lorsque l'auteur explique et justifie les procédures d'analyse [92] :

[92] **Nous avons donc choisi d'effectuer les estimations** à partir des postes forestiers de l'Igbp pour prendre en compte la part relative des différents écosystèmes présents par pays. Dans la « jungle des chiffres contradictoires », nous avons décidé de synthétiser et de ne retenir que des coefficients moyens, apparaissant acceptables au regard des travaux rencontrés. (geo.xml-s3299)

Dans cet exemple, le verbe *choisir* est lié à la narration scientifique et l'auteur présente sa démarche méthodologique. Cette fonction métatextuelle est plus marquée avec le temps du passé composé.

#### **b. Pronoms sujets mettant en jeu l'auteur**

De manière générale, l'emploi des pronoms sujets est différent pour les trois verbes étudiés. Dans la figure 4-20, deux cas se détachent dans le corpus : le verbe *choisir* préfère le pronom *nous* alors que les deux autres s'associent plus fréquemment au pronom *on*.



**Figure 4-20 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de choix)**

Si on examine l'association de *nous* et *choisir* de plus près, on constate que cette association n'apparaît que dans la construction *nous avons choisi de Vinf*. Comme nous l'avons indiqué plus haut, le verbe revêt ici une valeur métatextuelle et s'associe au sujet pronominal représenté par le *nous* auteur (singulier ou collectif).

Concernant les verbes *retenir* et *adopter*, le pronom *on* peut représenter essentiellement deux emplois selon le contexte : le *on* qui renvoie à l'auteur [93] ou le *on* qui implique l'auteur et la communauté scientifique [94].

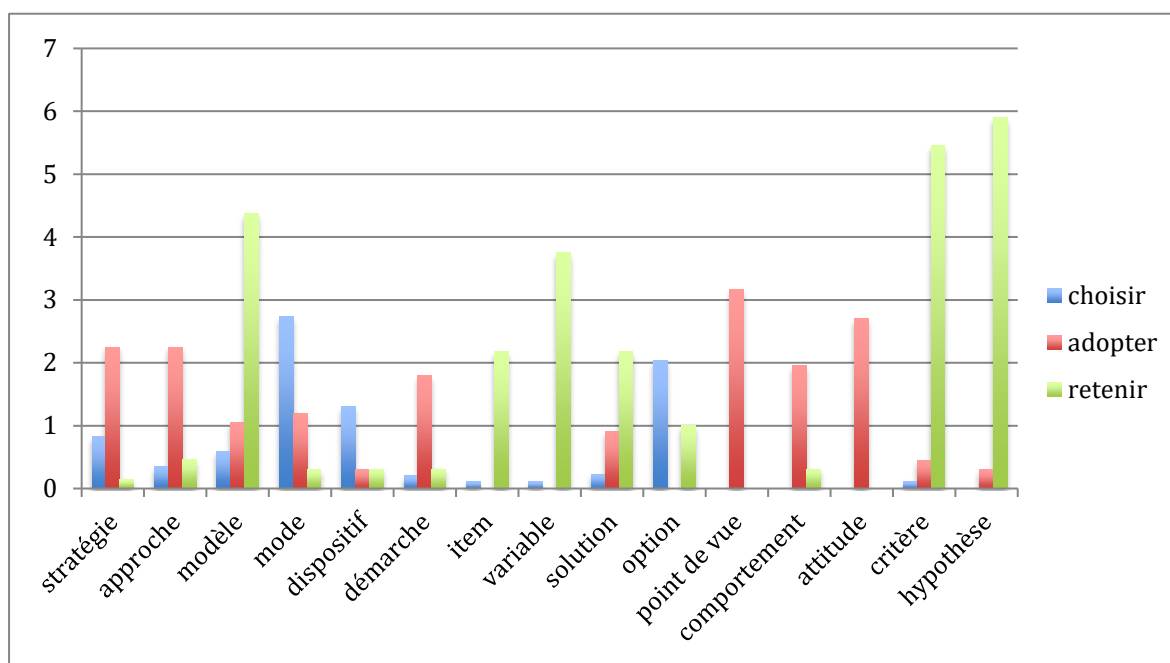
[93] **On a retenu les productions** les plus fréquentes afin de faciliter, lors de la présentation de la consigne, l'activation d'un stéréotype. (psycho.xml-s12026)

[94] Selon que l'**on adopte le point de vue** de Necka (1997) ou celui d'Eysenck et Calvo (1992), deux modèles explicatifs de l'effet de l'anxiété sont envisagés. Pour Necka, les stratégies interviennent pour donner une priorité soit aux mécanismes de stockage soit aux mécanismes de traitement. (psycho.xml-s13154)

### c. Arguments nominaux

L'examen de l'ensemble des objets noms les plus fréquents pour les trois verbes de choix permet de constater que ces noms renvoient majoritairement aux objets scientifiques. Notons par ailleurs que le verbe *choisir* reste moins productif que les

verbes *retenir* et *adopter*, comme l'illustre la figure 4-21.



**Figure 4-21 : Les noms objets des verbes de choix (en %)**

Les objets construits complexes et métascientifiques (*approche*, *modèle*, *mode*, *dispositif* et *démarche*) se rencontrent auprès des trois verbes, ce qui témoigne du transfert et de l'emprunt de ces objets d'un chercheur à l'autre. Ce constat rejoint l'étude de Grossmann *et al.* (2009), selon laquelle, les domaines SHS construisent des objets structurés « prêts à l'emploi » et il y a bien « un fonctionnement cumulatif des appareillages théoriques ou méthodologiques développés » (*ibid.*, p. 195).

Cependant, on trouve des différences au sujet du choix des objets nominaux relevant de la classe des objets scientifiques. Tout d'abord, les objets scientifiques autorisés par le verbe *adopter* peuvent renvoyer soit à la méthode de recherche scientifique (*démarche*, *approche*, *modèle*), soit à la pensée de la personne (*point de vue*, *comportement*, *attitude*). Par contre, il nous semble difficile de *choisir* ou *retenir* un *point de vue*. En effet, le verbe *adopter* est utilisé pour exprimer l'idée de se rallier à ce choix. Il a donc une dimension cognitive importante [95].

[95] Pour commenter nos résultats, nous **adopterons** d'abord **un point de vue** structuraliste puis, poussés par les faits, phénoménologique pour discuter de la spécificité d'une cognition spatiale âgée. (psycho.xml-s15184)

Ensuite, le verbe *retenir* privilégie les objets scientifiques du processus de raisonnement (*solution, hypothèse*) ou utilisés pour l'explication (*variable, critère, item*). Dans le dernier cas, il est difficilement remplacé par les deux autres verbes, car il implique le fait de conserver un élément après un examen critique, comme le montre l'exemple [96] :

- [96] Plusieurs autres variables de statut (revenus du ménage, nombre de subordonnés au travail) et d'opinion (opinion sur la démocratie, sur la concurrence ou sur la responsabilité des besoins individuels) ont été testées. Les premières n'ont pas été intégrées aux modèles présentés, dans la mesure où elles n'amélioreraient pas de façon significative leur qualité prédictive par rapport aux **variables** de contrôle déjà **retenues**. (sciencespolitiques.xml-s1202)

Globalement, au niveau syntaxique, cette sous-classe construit son objet de choix en autorisant l'alternance entre la construction avec complément nominal et le passif. L'objet scientifique a souvent tendance à être thématiqué. Au niveau sémantique, les verbes étudiés sont fortement associés à une justification de l'utilité instrumentale de l'objet scientifique choisi, à un effet provoqué par cet objet ou même à un objectif visé à l'aide de cet objet. Contrairement aux verbes d'opinion qui se caractérisent par un processus de modalisation, les verbes de choix sont davantage employés dans un contexte de raisonnement. Comme le soulignent Grossmann *et al.* (2009), le positionnement de l'auteur lié aux marques d'emprunt ou de choix reste relativement discret.

Les résultats de l'analyse sont indiqués dans les patrons ci-dessous :

choisir	<ol style="list-style-type: none"> <li>1. [[humain=auteur]] choisit de [v-inf=analyse] Ex : <i>Dans ce papier, après avoir fixé le nombre de facteurs à retenir à l'aide du test proposé, nous avons choisi de modéliser la dynamique des diverses composantes du modèle.</i> (eco.xml-s2242)</li> <li>2. [[humain=auteur]] choisit [[objet scientifique]] Ex : <i>Ce modèle a été choisi, car, conformément à notre cadre expérimental considérant la notion de signatures spatiales de la scène et de l'objet.</i> (psycho.xml-s10055)</li> </ol> <p>Lexset [[objet scientifique]] : <i>stratégie, approche, modèle, mode, dispositif, option</i></p> <p>Sens : Quelqu'un se détermine un objet scientifique entre deux ou plusieurs.</p>
---------	---

adopter	1. [[humain=auteur]] adopte [[méthode idée] Ex : <i>Face à ces difficultés, j'ai adopté une attitude pragmatique et exploratoire.</i> (anthropologie.xml-s13107) Lexset [[méthode idée]] : <i>stratégie, approche, modèle, mode, démarche, solution, point de vue, comportement, attitude</i> Sens : Quelqu'un choisit une méthode ou une idée et s'y rallie
retenir	1. [[humain=chercheur]] retient [[méthode élément explicatif thèse]] Ex : <i>Pour décompter les candidats dits de la « diversité », nous avons d'abord retenu deux critères.</i> (sciencespolitiques.xml-s11305) Lexset [[méthode élément explicatif]] : <i>modèle, item, variable, solution, option, critère, hypothèse</i> Sens : Quelqu'un conserve une méthode ou un élément après l'examen critique
Valeurs sémantiques	valeur évaluative (s'associer au raisonnement causal qui peut mettre en place une justification, un effet ou un objectif)
Routines	<i>en choisissant</i> (25/837) <i>en adoptant</i> (28/664) <i>en retenant</i> (11/640)
Commentaires	Le passif et la construction relative sont prédominants, le gérondif est bien présent, le verbe <i>choisir</i> s'associe souvent au pronom <i>nous</i> .

**Tableau 4.12 : Patrons lexico-syntaxiques des verbes de choix**

#### 4.4.5. Verbes d'hypothèse

La formulation d'hypothèses en réponse à un problème de recherche est considérée comme la « base intellectuelle » de l'écrit scientifique (Lindsay & Poindron, 2011, p. 33). Les hypothèses énoncées permettent de guider les objectifs de l'étude. L'auteur doit « faire l'effort de prédire ce qui pourrait être logiquement attendu comme résultat » (*ibid.*, p. 33) et l'objectif de l'étude consiste à confirmer ou rejeter ces hypothèses, amenant ainsi vers la réalisation d'un fait ou des conséquences. La formulation d'hypothèses peut se réaliser à l'aide des collocations telles que *faire une hypothèse, émettre une hypothèse, poser une hypothèse*. Concernant les verbes d'hypothèse qui exprime le sens de 'poser quelque chose comme hypothèse servant de base à un raisonnement, à une argumentation' (d'après le TLFi), nous avons retenu deux verbes *supposer* (1062) et *postuler* (130). Nous nous concentrons maintenant sur l'analyse du comportement sémantique et syntaxique de ces deux verbes permettant d'introduire une hypothèse.

##### a. Constructions syntaxiques et valeurs sémantiques

Pour les deux verbes étudiés, les structures privilégiées dans la formulation d'hypothèses sont la construction en complétive *que* [97] et la construction avec



complément nominal [98], moins fréquente. Pour ces deux constructions, on a constaté une tendance de modalisation avec *pouvoir* (*pouvoir supposer* 99/1062, *pouvoir postuler* 16/130) :

[97] En effet, **nous pouvons supposer que**, malgré l'effort de mémorisation, ce type d'apprentissage n'est plus opérationnel dès lors que l'élève est confronté à une épreuve qui implique la réutilisation des savoirs appris dans un contexte sensiblement différent. (scienceseducation.xml-s1671)

[98] Néanmoins, la motivation extrinsèque ne doit pas être uniquement considérée comme une motivation non autodéterminée dans la mesure où des individus peuvent être extrinsèquement motivés et se sentir autonomes (Deci et Ryan, 2008). Ainsi, Deci et Ryan (1985) ont **postulé l'existence** de quatre formes de motivation extrinsèque : la régulation externe, la régulation introjectée, la régulation identifiée et la régulation intégrée. (psycho.xml-s39)

La proposition complétive ainsi que le complément nominal introduit sont considérés comme des faits vrais ou probables. Ici, les verbes sont employés dans un raisonnement logique puisque l'hypothèse est souvent accompagnée d'une justification, comme en témoigne la présence non négligeable des adverbes discursifs (*cependant, donc, alors, en effet, par ailleurs, en revanche, ainsi*) qui justifient l'introduction de l'hypothèse [99] :

[99] Notre but en choisissant la presse était de rencontrer des figures présumées non littéraires, et donc de pouvoir considérer sous un angle différent ce qui est appelé généralement " figure de style ". **Nous supposons en effet que**, dans un discours informatif tel que celui de la presse, l'objectif de l'énonciateur est centré sur le contenu à transmettre au moins autant que sur la forme du message... (ling.xml-s2824)

Dans l'exemple ci-dessus, l'auteur énonce son hypothèse suite à l'explication de l'objectif de son étude, permettant au lecteur de suivre son raisonnement.

En fait, le verbe *supposer* est plus productif en termes de constructions syntaxiques : la structure hypothétique à l'impératif (50/1062) et des structures causatives telles que *laisser supposer que* (15/1062), *conduire/amener à supposer* (10/1062). Il apparaît souvent concomitamment dans un co-texte avec l'expression de cause ou de

conséquence. Les deux exemples suivants illustrent cette association :

[100] **Supposons que** les bilingues, conscients qu'ils doivent traiter successivement deux blocs de langues différentes, **puissent** d'une certaine manière adapter a priori l'activation de leurs deux langues. Dans ce cas, les résultats obtenus pourraient refléter une combinaison de processus lexicaux (compétition entre les différents candidats activés) et extra-lexicaux (activation a priori des deux langues). (psycho.xml-s9630)

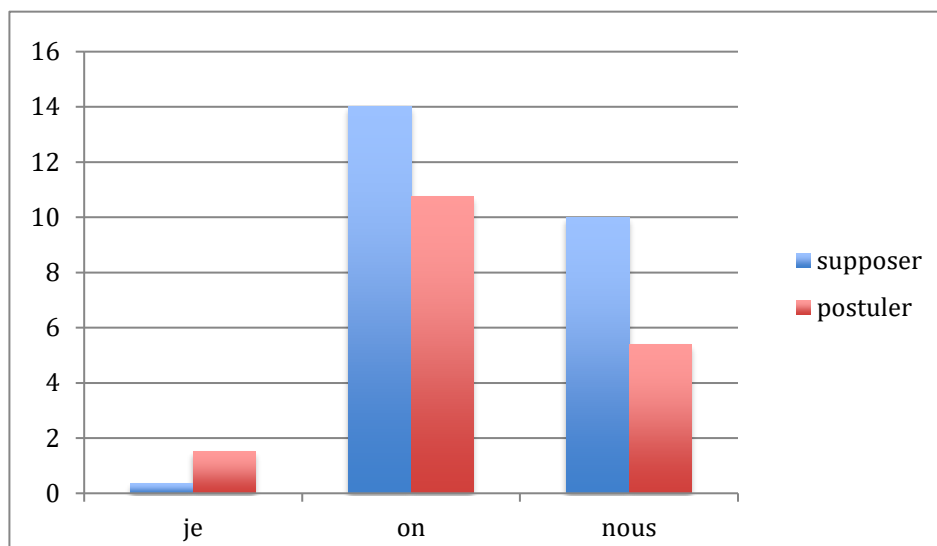
[101] Ce constat métathéorique souligne, s'il en était besoin, l'intérêt et la portée des résultats obtenus dans cette recherche. C'est ainsi qu'on table volontiers sur la persuasion pour inciter autrui à modifier ses idées ou ses conduites. **Nos résultats laissent supposer que** lorsqu'on veut obtenir un comportement particulier de la part de quelqu'un, il est pertinent de créer les circonstances qui l'amèneront à décider par lui-même de le réaliser. (psycho.xml-s15540)

Dans l'exemple [100], le verbe *supposer* à l'impératif introduit un fait que l'auteur considère comme probable. Cette supposition permet ensuite de déduire les résultats éventuels. Quant à l'exemple [101], la fonction de la construction causative *nos résultats laissent supposer que* est d'introduire une interprétation et une hypothèse à partir des résultats obtenus. Ici, l'auteur justifie le bien-fondé de son hypothèse, ce qui favorise la crédibilité de son discours.

Au niveau de constructions syntaxiques, il apparaît que cette sous-classe a des points communs avec les verbes d'opinion. En effet, ces verbes se caractérisent par une tendance à être modalisés et mettent en avant un raisonnement bien fondé.

#### **b. Pronoms sujets mettant en jeu l'auteur**

De même que les verbes d'opinion, les deux verbes d'hypothèse marquent une fréquence relativement élevée du pronom *on*, comme l'illustre la figure 4-22 :



**Figure 4-22 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes d’hypothèse)**

La forme modalisée *pouvoir* est privilégiée pour le pronom *on*. Notons par ailleurs qu’elle apparaît davantage dans la partie du développement des articles du corpus, servant à introduire une interprétation atténuée ou à poser une affirmation nuancée [102] :

[102] Dans les cas de maladie d’Alzheimer, on n’observe pas cette « aide » exceptionnelle de voisinage. **On peut supposer que** la perte d’autonomie (au sens fort, c’est-à-dire l’aliénation mentale, alors que la personne peut avoir conservé son indépendance physique) décourage de telles restitutions... (sociologie.xml-s11493)

Nous avons vu cette association (*on peut l’interpréter comme, on peut considérer que*) est liée à une fonction de négociation. Ici, le pronom *on* semble impliquer la communauté de discours et le lecteur dans son interprétation personnelle et le verbe prend une valeur interprétative.

Par ailleurs, nous avons également relevé une présence non négligeable du pronom du pronom *nous*. Il peut renvoyer soit à l’auteur unique (singulier ou collectif) [103] soit à l’auteur et à la communauté de discours [104] :

[103] L’objectif de cet article étant de caractériser les effets engendrés par une pratique de certification, volontaire et coûteuse, sur le fonctionnement d’un tel marché, **nous supposons que** les firmes ont la possibilité de se signaler auprès des consommateurs

par l'adoption d'un audit ou test de la qualité, volontaire et coûteux. (économie\_229)

[104] En effet, **nous pouvons supposer que**, malgré l'effort de mémorisation, ce type d'apprentissage n'est plus opérationnel dès lors que l'élève est confronté à une épreuve qui implique la réutilisation des savoirs appris dans un contexte sensiblement différent  
7. (scienceseducation.xml-s1671)

Dans l'exemple [104], le verbe introduit une hypothèse qui doit être confirmée ultérieurement, prenant ainsi la valeur hypothétique. De manière tendancielle, il semble que dans l'introduction, concernant la formulation des hypothèses, cette combinaison *nous* auteur+*supposer/postuler* est fréquente, à côté de l'objectif de l'étude et des questions posées.

### c. Arguments nominaux

Contrairement aux verbes d'opinion, les sujets métonymiques ne sont pas exclus avec les verbes d'hypothèse, mais ils sont peu fréquents.

Au niveau de l'objet, les noms sont peu productifs (en particulier pour *supposer*). Ceci peut être lié au faible nombre de constructions avec complément nominal. De manière tendancielle, les objets nominaux autorisés par les deux verbes sont des noms abstraits tels que *différence*, *existence*, *relation*, *nécessité*, etc. Le faible nombre des objets nominaux ne nous permet pas de proposer une interprétation au niveau sémantique, mais il est intéressant de constater la présence du raisonnement causal, comme le montrent les exemples [105] et [106] :

[105] Il nous semble donc raisonnable de nous aligner sur les tendances de la littérature **en postulant des liens** positifs lorsqu'ils sont significatifs. (psycho.xml-s1601)

[106] Lorsque nous parlons de compréhension, nous sommes amenés à **supposer l'existence** d'un savoir ordinaire des uns et des autres, savoir qui repose sur des typifications et des attributions. (sociologie.xml-s8004)

Dans ces exemples, l'hypothèse posée se rapporte à la relation entre deux objets (*lien*) ou à l'état d'un objet d'étude (*existence*).

À l'issue de cette analyse, nous avons montré que *supposer* et *postuler* peuvent

être utilisés non seulement pour introduire une hypothèse en amont de l'étude, mais aussi marquer une interprétation des résultats. Dans l'expression de l'hypothèse, ces verbes mobilisent préférentiellement la construction en complétive *que*. En outre, ces verbes sont à rapprocher des verbes d'opinion en ce qu'ils mettent l'accent sur le processus de modalisation et de raisonnement, et l'on observe la fréquence élevée du modalisateur *pouvoir*, des adverbes discursifs et des constructions causatives.

Les patrons lexico-syntaxiques de ces deux verbes sont présentés dans le tableau 4-13 :

supposer	<p>1. [[humain=auteur, chercheur]] (pouvoir) [adverbe discursif] supposer [complétive <i>que</i>] Ex : <b>On suppose que</b> la concurrence est pure et parfaite sur le marché des biens... (eco.xml-s1055)</p> <p>2. [[humain=auteur, chercheur]] (pouvoir) [adverbe discursif] supposer [[relation]propriété] Ex : Dans la discussion, <b>nous supposons l'existence</b> de modes de gestion de l'espace ayant des répercussions potentielles sur la capacité des enfants à engager des relations sociales. (scienceseducation.xml-s5427)</p> <p>Lexset : <i>existence, nécessité, lien, différence, relation, rapport</i> Sens : Quelqu'un pose quelque chose comme hypothèse servant de base à un raisonnement, à une argumentation. Ici, l'expression de l'hypothèse est souvent accompagnée de la justification et de la conséquence.</p>
postuler	<p>1. [[humain=auteur, chercheur]] (pouvoir) [adverbe discursif] postuler [complétive <i>que</i>] Ex : Dans le cas de l'écriture, <b>on peut postuler qu'il en va de même</b> entre le concept et le graphème. (anthropologie.xml-s10996)</p> <p>2. [[humain=auteur, chercheur]] (pouvoir) [adverbe discursif] postuler [[relation]propriété] Ex : Il nous semble donc raisonnable de nous aligner sur les tendances de la littérature en <b>postulant des liens positifs</b> lorsqu'ils sont significatifs. (psycho.xml-s1601)</p> <p>Lexset [[relation]propriété] : <i>existence, lien, affinité, fonction, nécessité, relation, rapport</i> Sens : Quelqu'un admet (à titre d'hypothèse) quelque chose comme vrai afin d'en déduire un fait.</p>
Valeurs sémantiques	valeur interprétative, valeur hypothétique (s'associer au raisonnement causal qui peut mettre en place une justification, un effet ou un objectif)
Routines	<i>supposons que</i> (50/1062) <i>laisser supposer que</i> (15/1062) <i>conduire/amener à supposer que</i> (10/1062) <i>en postulant</i> (7/130)/ <i>en supposant que</i> (54/1062)
Commentaires	le sujet est souvent <i>nous</i> et <i>on</i> , sujets métonymiques possibles ( <i>hypothèse, théorie, approche, travail, analyse</i> )

**Tableau 4.13 : Patrons lexico-syntaxiques des verbes d'hypothèse**

#### 4.4.6. Verbes de proposition

La proposition scientifique consiste en un apport spécifique de l'auteur (Tutin, 2010a). Elle relie l'auteur (rôle de chercheur) aux objets construits par l'activité scientifique (*méthode, approche*) ou aux procédés d'analyse (*définition, interprétation*). En effet, la proposition d'une approche nouvelle ou d'une analyse différente contribue à mettre en relief la valeur distinctive ou la contribution de l'étude. Il existe donc un enjeu de positionnement pour situer son apport par rapport aux pairs de la communauté scientifique (Grossmann *et al.*, 2009). Deux verbes représentatifs *proposer* (2247) et *suggérer* (520) sont relevés. Dans le LVF, ces verbes sont considérés comme verbes de communication et prennent le sens de 'communiquer un conseil'. Cette définition montre clairement que la dimension communicative est importante, et dans le contexte scientifique, l'objet indirect désigne le lecteur, récepteur de la proposition. Cette sous-classe implique donc non seulement la prise de position, mais aussi l'acte communicatif par lequel l'auteur présente quelque chose à l'examen, à la réflexion ou à l'évaluation du lecteur.

##### a. Constructions syntaxiques et valeurs sémantiques

Concernant les constructions en commun, ces deux verbes prennent un sujet humain et un objet inanimé. Il faut noter que cette construction est notamment fréquente pour le verbe *proposer* [107].

[107] **Ducrot (1984) propose une conception** polyphonique de l'interrogation rhétorique où le questionneur fait "entendre" la voix de l'allocutaire se posant cette question. (ling.xml-s10072)

Par ailleurs, comme en témoignent les exemples [108] et [109], ces verbes peuvent autoriser l'alternance du passif ou du passif réduit, le verbe *proposer* restant beaucoup plus productif.

[108] À la différence des études qui se focalisent sur un type d'accident particulier (exemple des accidents à un véhicule et un piéton), **la méthode proposée ici** permet de prendre en compte tous les types d'accidents dans une même analyse. (psycho.xml-s4214)

[109] Comme dans l'étude de Detterman et Daniel (1989), une analyse confirmatoire "multigroupes" des matrices de corrélations a été conduite à l'aide de la **méthode suggérée** par Werts, Rock, Linn et Jöreskog (1976). (psyco\_35)

Dans les deux exemples ci-dessus, l'objet proposé (*méthode*) est mis en avant. Il s'agit d'introduire une notion, une méthode nouvelle ou empruntée à d'autres chercheurs qui correspondent ici aux compléments d'agent. Il apparaît que l'important est d'attirer l'attention du lecteur sur les objets scientifiques et non de parler de la personne qui a proposé ces objets. Il est à noter que dans cet emploi, le verbe *proposer* est fréquemment associé au déictique de lieu *ici* (46/2247) qui signifie 'sur ce sujet'. Ce déictique permet à l'auteur de mettre l'accent sur son apport particulier [108].

En dehors de ces constructions, les deux verbes peuvent également être utilisés pour présenter les activités d'analyse lorsqu'ils acceptent un verbe à infinitif introduit par la proposition *de* (*proposer de Vinf*: 234/2247, *suggérer de Vinf*: 19/520) comme le montrent les exemples [110] et [111] :

[110] Pour clarifier ce point, **je propose de comparer** ce mode de classement des animaux, seulement suggéré par les textes, avec un autre qui, lui, est tout à fait explicite : leur partition en purs et impurs. (anthropologie\_36)

[111] Face à ce constat, **Ackerman [1997] suggère de développer** l'approche consommateur producteur afin (i) de comprendre la consommation comme activité socialisée reflétant certaines interactions, (ii) de considérer les désirs de consommation (wants) selon leur urgence relative et leur possibilité de saturation. (économie\_267)

Dans l'exemple [110], l'auteur présente son activité d'analyse à la réflexion et à l'évaluation du lecteur. Il est important de souligner que les verbes à infinitif autorisés appartiennent souvent aux verbes d'analyse tels que *étudier*, *analyser*, *vérifier*, *distinguer*, etc. Il s'agit de présenter la procédure de l'analyse comme devant être effectuée. Cet emploi rapproche ces deux verbes du verbe *choisir* que nous avons présenté plus haut.

De plus, le gérondif a également été constaté pour ces deux verbes (*en proposant* 57/2247, *en suggérant* 15/520). Ici, il peut introduire la manière et la simultanéité

[112] :

[112] Nous confirmons cette conclusion **en proposant un exemple** simple avec préférences CES et technologie Cobb - Douglas dans lequel l'équilibre stationnaire est localement indéterminé lorsque la consommation des agents à durée de vie infinie est suffisamment faible. (eco.xml-s14077)

Au-delà de ces constructions communes, ces deux verbes ont leurs propres caractéristiques au niveau syntaxique. En fait, le verbe *suggérer* se distingue par le fait qu'il privilège de manière assez nette la construction en complétive *que* [113] :

[113] Le principal argument repose sur la sensibilité du test formel à la spécification retenue. Ainsi, **TAYLOR [1992] suggère que** des imperfections "mineures" sur les données pourraient conduire au rejet de la théorie, alors même que les deux pentes des taux seraient très proches. (eco.xml-s10293)

Dans l'exemple ci-dessus, le verbe *suggérer* introduit l'assertion d'un autre chercheur qui sert à l'argument de l'auteur. Ainsi, ce verbe est à rapprocher des autres verbes qui remplissent la fonction de renvoi (*expliquer, montrer, affirmer, souligner*). Cette fonction est également associée aux constructions en *comme* (44/520) : *comme le suggère, comme suggéré*, voir l'exemple [114] :

[114] De 1476 à 1500, plus de 30 % de la noblesse flamande se composait de familles d'origine extrarégionale. Il est donc vrai qu'au cours du bas Moyen Âge, la noblesse de Flandre et celle d'autres régions se sont intimement mêlées et de façon grandissante, **comme l'ont suggéré** Henri Pirenne et un grand nombre d'autres historiens. (histoire.xml-s11576)

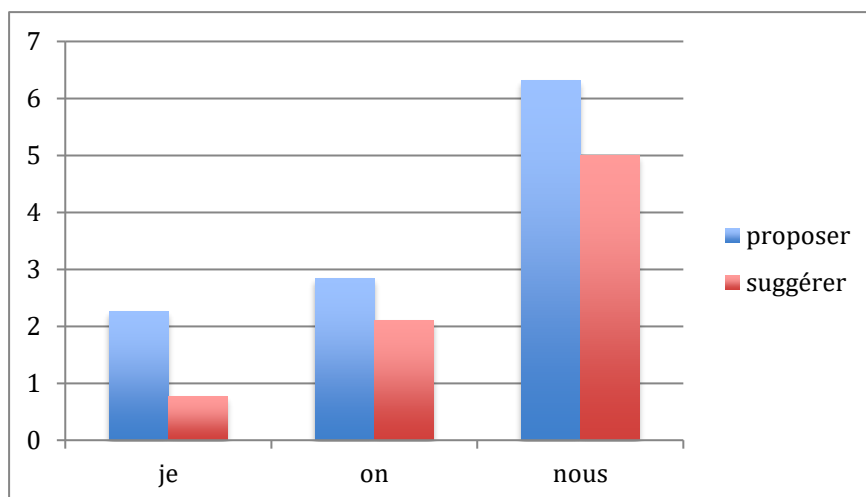
Notons d'ailleurs que la construction en *comme* reste marginal pour le verbe *proposer* (15/2247).

À la différence de *suggérer*, *proposer* peut s'employer dans des constructions exprimant l'objectif ou introduisant le thème abordé, avec par conséquent une valeur métatextuelle importante : *l'objectif est de proposer* (7/2247), *il s'agit de proposer* (8/2247), *essayer de proposer* (5/2247), *consister à proposer* (5/2247).



## b. Pronoms sujets mettant en jeu l'auteur

Parmi les trois pronoms examinés, le pronom *nous* est le plus fréquent (voir la figure 4-23) :



**Figure 4-23 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes de proposition)**

Ici, il s'agit du pronom *nous* auteur (singulier ou collectif). En effet, dans la proposition scientifique, l'emploi exclusif du *nous* a pour fonction principale de souligner l'apport scientifique propre à l'auteur ou sa prise de position. Considérons les exemples suivants :

[115] Bien que les résultats de Major et *al.* (1993) continuent d'être fréquemment cités et utilisés pour développer certaines propositions théoriques, des travaux plus récents tendent à montrer que... Ainsi, **nous suggérons que** les résultats obtenus par Major et *al.* (1993) ne s'appliqueraient pas dans un contexte qui menace l'identité sociale. (psycho.xml-s12387)

[116] **Nous voudrions, dans les pages qui suivent, proposer une réflexion** autour de l'emploi de termes comme « travestissement » et « imposture » à propos d'un cas tout à fait spécifique : la petite centaine (chiffre incertain) de femmes soldats recensées pendant la Révolution française et l'Empire [2]. (sciencespolitiques.xml-s22183)

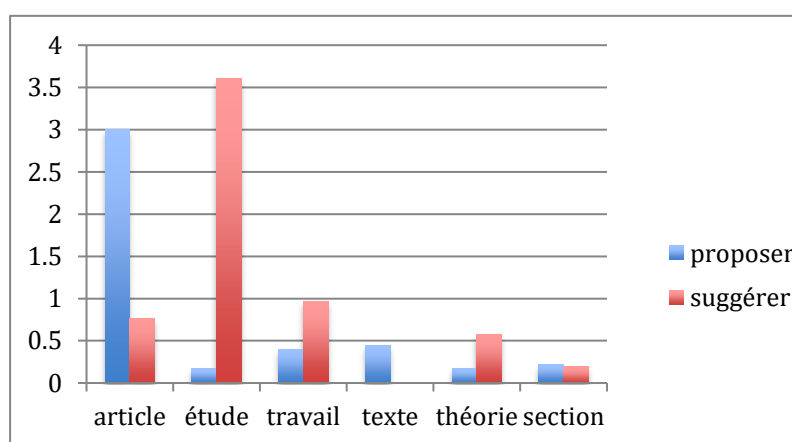
Dans l'exemple [115], l'auteur affiche son idée par rapport aux pairs. Quant à l'exemple [116], l'auteur souligne sa contribution (*réflexion*) à la connaissance

scientifique.

Outre l'emploi de ces pronoms, la proposition de l'auteur peut également se réaliser à travers l'emploi métonymique.

### c. Arguments nominaux

La figure 4-24 ci-dessous illustre la répartition des sujets métonymiques les plus fréquents pour cette sous-classe. Le verbe *proposer* s'associe fréquemment au nom *article* tandis que *suggérer* préfère le nom *étude*.



**Figure 4-24 : Les sujets métonymiques des verbes de proposition (en %)**

Si l'on examine l'emploi des sujets *article* et *étude* de plus près, on constate que l'expression *cet article propose* apparaît souvent dans le résumé ou l'introduction, revêtant ainsi une valeur métatextuelle. Par exemple, dans l'exemple [117], le verbe *proposer* sert à introduire l'objectif de l'étude :

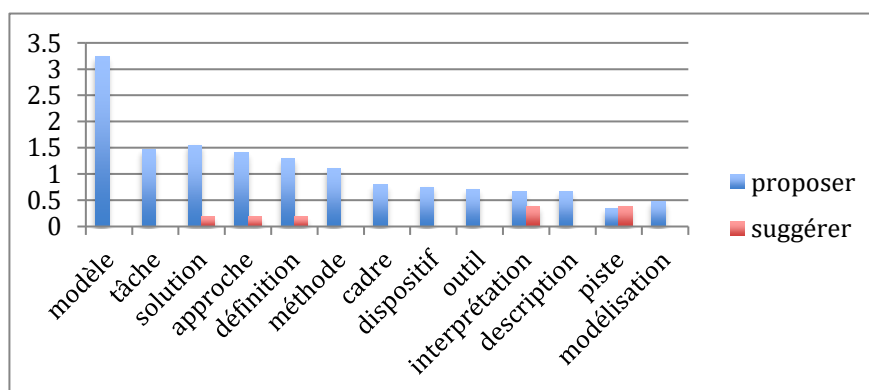
[117] **Cet article propose** une analyse empirique des transformations de l'espace régional européen du point de vue de l'agglomération des activités économiques et de recherche, et des spécialisations sectorielles dans les régions dites « centrales » et « périphériques ». (sociologie.xml-s16190)

Contrairement au verbe *proposer* qui se rapporte à l'article de l'auteur, le verbe *suggérer* préfère le nom *étude* comme sujet métonymique. D'une part, ceci peut renvoyer à l'étude de l'auteur, comme dans l'exemple [118], où l'auteur met en avant son opinion en la soumettant à la réflexion du lecteur. D'autre part, il peut renvoyer aux travaux d'autrui qui servent les arguments de la recherche [119] :

[118] **Notre étude suggère que** les enfants utilisent des modèles d'indices graphiques typiques pour marquer les émotions. Ainsi, l'homme joyeux est représenté au moyen d'une bouche concave, de bras levés et d'un contexte de type paysage. (psycho.xml-s14428)

[119] **Plusieurs études récentes suggèrent** d'autre part **que** des rapports de patronage importants pouvaient être établis par la parenté spirituelle : qu'en est-il dans le cas des Morend ? (histoire.xml-s21948)

Au niveau de l'objet, le verbe *proposer* est beaucoup plus productif (voir la figure 4-25). Ce résultat est lié au fait qu'il privilégie la construction avec un objet nominal.



**Figure 4-25 : Les noms objets des verbes de proposition (en %)**

Deux classes sémantiques peuvent être distinguées : d'une part, les objets nominaux peuvent être des noms renvoyant aux objets construits de l'activité scientifique (*modèle, approche, solution, piste*), en particulier ceux qui relèvent de la méthode de recherche scientifique tels que *méthode, méthodologie, dispositif, outil*, etc. [120] ; d'autre part, les objets nominaux peuvent marquer un processus cognitif comme *interprétation, description, modélisation et définition* [121] :

[120] Ici, nous **proposons une méthode** permettant de tester l'hypothèse de marchés complets en tenant compte de l'hétérogénéité des préférences et, en particulier, des aversions au risque. (eco.xml-s15876)

[121] Cet article conjugue perspectives théoriques et appliquées dont la validité scientifique paraît établie, et il fait l'objet de nombreuses citations dans la discipline. L'auteur y

**propose une définition** des CRS en tant que compétences stratégiques.

scinfo.xml-s11042

Le faible nombre des objets nominaux autorisés par le verbe *suggérer* ne nous permet pas de proposer une généralisation.

Enfin, l'analyse nous permet de constater que les verbes de proposition expriment non seulement l'opinion (*nous suggérons que*) et l'intention (*nous proposons d'étudier*) de l'auteur, mais aussi l'apport scientifique (*méthode que nous proposons*). Au sein de cette sous-classe, *proposer* et *suggérer* présentent des traits différents tant au niveau sémantique que syntaxique : le premier est davantage mobilisé pour mettre en avant l'intention et l'apport scientifique, ayant parfois une valeur métatextuelle alors que le deuxième fréquemment utilisé dans la citation revêt une valeur argumentative.

Nous présentons maintenant un récapitulatif des patrons verbaux dans le tableau ci-dessous :

proposer	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] propose [[objet scientifique processus cognitif]] Ex : <i>Pour examiner cette question, nous proposons l'analyse d'un article de référence en gestion, celui de Persais (2004).</i> (scinfo.xml-s11040)</li> <li>2. [[humain=auteur, chercheur]] propose de [v-inf=analyse] Ex : <i>Dans une perspective modélisatrice, nous proposons d'en étudier, sous le nom d'économie symbolique, un modèle extrêmement simplifié...</i> (anthropologie.xml-s10597)</li> </ol> <p>Lexset [[objet scientifique processus cognitif]] : <i>méthode, approche, interprétation, définition, outil, dispositif, modèle</i> Sens : Quelqu'un présente un objet scientifique, un processus cognitif ou une activité scientifique (qui sera intégré dans la recherche) à la réflexion et à l'évaluation du lecteur.</p>
suggérer	<ol style="list-style-type: none"> <li>1. [[humain=auteur, chercheur]] suggère [[raisonnement processus cognitif thèse]] Ex : <i>L'article suggère quelques pistes qui permettent d'éviter les explications ad hoc.</i> (sociologie.xml-s3903)</li> <li>2. [[humain=auteur, chercheur]] suggère de [v-inf=analyse] Ex : <i>Les sociologues mathématiciens suggèrent d'élever le niveau de scientificité de la sociologie.</i> (sociologie.xml-s5075)</li> <li>3. [[humain=auteur, chercheur]] suggère [complétive que] Ex : <i>Ainsi, TAYLOR [1992] suggère que des imperfections "mineures" sur les données pourraient conduire au rejet de la théorie...</i> (eco.xml-s10293)</li> <li>4. Comme (le) suggère [[humain=auteur, chercheur]] Ex : <i>Comme suggéré plus haut, le premier volume de l'Histoire de la sexualité (1976) paraît prolonger cette perspective.</i> (scienceseducation.xml-s10146)</li> </ol> <p>Lexset [[raisonnement processus cognitif thèse]] : <i>approche, définition, solution, piste, interprétation</i></p>

	Sens : Quelqu'un présente une idée au lecteur pour le faire accepter.
Valeurs sémantiques	valeur métatextuelle, valeur argumentative
Routines	<i>en proposant</i> 57/2247 <i>en suggérant</i> 15/520 <i>l'objectif est de proposer</i> (7/2247) <i>il s'agit de proposer</i> (8/2247) <i>essayer de proposer</i> (5/2247) <i>consister à proposer</i> (5/2247)
Commentaires	Le pronom <i>nous</i> est plus fréquent. Sujets métonymiques possibles ( <i>article propose, étude suggère</i> ), le passif et la construction relative sont fréquents pour le verbe <i>proposer</i>

**Tableau 4.14 : Patrons lexico-syntaxiques des verbes de proposition**

#### 4.4.7. Verbes d'inférence

De même que l'activité d'interprétation, l'inférence implique un raisonnement rigoureux où l'auteur tire une conséquence et établit un fait scientifique à partir d'un ensemble d'éléments (données, analyse, etc.). Nous avons relevé deux verbes d'inférence tels que *conclure* (392) et *déduire* (176), interprétés par le LVF comme des verbes qui ont le sens commun de 'faire sortir un fait ou un résultat d'un élément abstrait'. Autrement dit, ces deux verbes sont trivalents. Dans l'inférence, l'agent A déduit un élément B d'un autre élément C, C étant ainsi la source de B, comme l'illustre la formule suivante :

N [agent A] +V+ de N [source B] + N [résultat C]

L'élément C peut également être introduit par une complétive *que*.

N [agent A] +V+ de N [source B] + que [résultat C]

Notons par ailleurs qu'en plus des verbes étudiés, l'inférence peut aussi être introduite par des verbes de relation (*il ressort de l'analyse que, il en découle/résulte que*) ou des connecteurs (*par conséquent, en conséquence*).

##### a. Constructions syntaxiques et valeurs sémantiques

Deux constructions principales ont été relevées. La première est celle de la complétive *que* et il arrive souvent que la construction contienne le pronom

anaphorique *en*, comme l'illustre les exemples [122] et [123] :

[122] À l'inverse, on observe que, parmi les utilisateurs occasionnels d'Internet, aucun n'a d'amis virtuels. **On peut en déduire que** l'utilisation d'Internet comme média relationnel ne se fera qu'une fois l'internaute familiarisé avec Internet, cette familiarisation passant par des usages d'Internet qui ne répondent pas nécessairement à un besoin de socialisation. (scinfo.xml-s4478)

[123] Quant au reste de la noblesse, celle-ci ne comptait alors que 10,9 % de nobles allochtones. **On peut en conclure que**, pour la haute noblesse, l'existence des réseaux interrégionaux a certainement précédé de beaucoup la naissance de l'État [50]. (histoire.xml-s11722)

Dans les deux exemples ci-dessus, le pronom *en* renvoie à l'observation [122] et aux données [123] sur lesquelles se fonde l'inférence des auteurs.

Ensuite, le verbe *déduire* peut s'associer à la construction nominale qui marque le résultat [124] :

[124] Certaines modalités peuvent donc être associées, avec un très faible taux d'erreur, à l'absence d'amis virtuels, et **on peut donc en déduire** un profil social qui se traduit dans 95 % des cas par l'absence d'amis virtuels. (scinfo.xml-s4484)

Notons également que ces deux constructions principales mobilisent souvent le modalisateur *pouvoir* (*pouvoir conclure* 52/392, *pouvoir déduire* 30/176) ainsi que des adverbes de déduction tels que *ainsi* et *donc*. Comme expliqué précédemment, ces marqueurs linguistiques mettent en avant une dimension argumentative. Dans l'exemple [124], le *pouvoir* associé au pronom *on* ouvre une négociation où l'auteur implique la communauté de discours et le lecteur dans le processus d'inférence. D'ailleurs, l'adverbe *donc* marque la conséquence, mettant en relief un raisonnement causal.

Comme indiqué plus haut, ces deux verbes sont étroitement liés au raisonnement causal, ce qui est renforcé par le fait qu'ils établissent une inférence à l'aide des constructions causatives : *permettre de conclure* (21/392), *conduire à conclure* (6/392), *amener à conclure* (5/392), *permettre de déduire* (7/176). On a également

relevé des constructions avec le verbe impersonnel introducteur *falloir* : *il faut conclure que* (11/392), *il faut en déduire que* (5/176), ce qui marque un effacement énonciatif. Ainsi, on peut rapprocher ces verbes des verbes d'opinion.

À la différence du verbe *conclure*, *déduire* peut être utilisé au passif [125] ou au pronominal passif [126] pour mettre en relation la source et le résultat de l'inférence :

[125] D'autres conclusions **peuvent être déduites** du dénombrement, qu'elles concernent les mots (sachant que les mots les plus courants en français ont 1,2 syllabe en moyenne, soit trois à quatre phonèmes) ou les phrases produites journallement. (ling.xml-s1785)

[126] Le matériau ligneux **peut se déduire** du fait que les dessins ne montrent ni motif sculpté, ni renflement aux extrémités, à la différence des caducées des vases à figures rouges postérieurs. (histoire.xml-s28235)

À l'issue de cet examen des constructions syntaxiques, il semble que les verbes *déduire* et *conclure* mettent en jeu un processus de l'effacement énonciatif. Nous allons maintenant examiner leur association avec les trois pronoms sujets.

### b. Pronoms sujets mettant en jeu l'auteur

En observant de plus près les pronoms sujets mettant en jeu l'auteur, on s'aperçoit que le pronom *on* est de loin le plus fréquent, en particulier pour le verbe *déduire*, comme le montre la figure 4-26 :

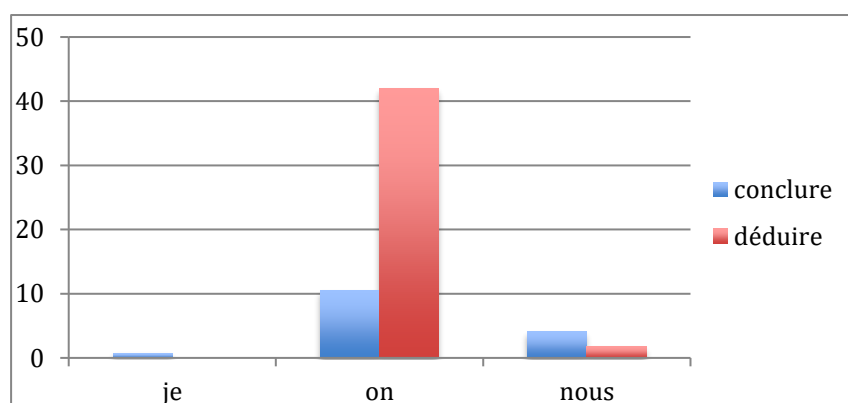


Figure 4-26 : Les fréquences relatives des pronoms sujets *je*, *on* et *nous* (en % en co-occurrences des verbes d'inférence)

L'emploi du pronom *on* avec le modalisateur *pouvoir* est dominant. Il a été

analysé précédemment dans l'étude des verbes d'opinion et des verbes d'hypothèse. En effet, en incluant la communauté de discours et le lecteur dans son argumentation, l'auteur les persuade de la validité de son inférence, la conclusion obtenue étant ainsi convaincante [127] :

[127] Soit l'expression "j'ai mis le vin dans le frigo". Elle fait référence à l'objet déplaçable et localisable, une bouteille par exemple, contenant du vin. À partir de cette interprétation, **on peut déduire que** le lexème vin peut donner accès à des informations concernant le contenant du vin. (ling.xml-s7752)

L'emploi du *nous* reste marginal, mais il peut également représenter l'auteur, la communauté de discours et le lecteur, cet emploi est renforcé par le modalisateur *pouvoir* (Grossmann & Tutin, 2010b, p. 8), comme le montre l'exemple [128] :

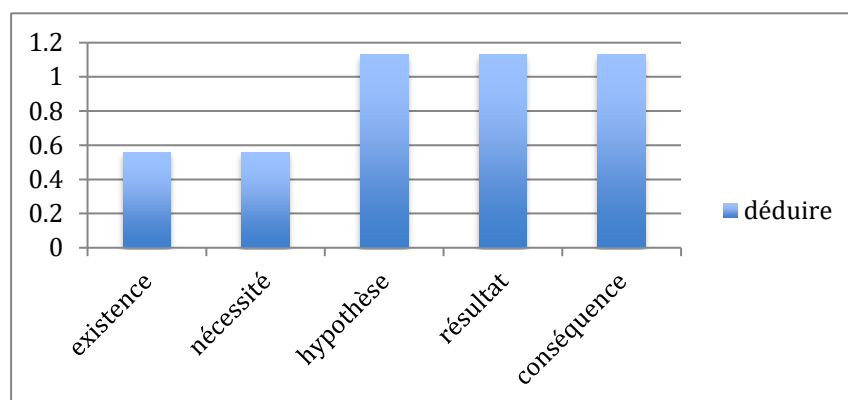
[128] **Nous pouvons conclure que** des comportements électoraux propres aux minorités ethniques sont autant présents en Belgique qu'aux Pays-Bas [42], malgré les structures des opportunités discursives défavorables pour une mobilisation ethnique à Bruxelles. **En effet**, nos résultats montrent que les électeurs d'origine non européenne présentent un comportement électoral significativement différent de celui des « Belgo – Belges ». (sciencespolitiques.xml-s21218)

Dans l'exemple, cette prise de position à partager est accompagnée d'une justification. La proposition comportant *en effet* sert à justifier la conclusion précédente. De ce fait, la dimension argumentative est bien présente.

### c. Arguments nominaux

Comme nous l'avons déjà souligné, contrairement au verbe *conclure* qui est employé préférentiellement avec une complétive *que*, le verbe *déduire* peut s'associer à la construction nominale qui marque le résultat. Nous nous intéressons ici au complément d'objet direct puisque le complément d'objet indirect est souvent remplacé par le pronom anaphorique *en*. La figure 4-27 donne un aperçu de la fréquence.





**Figure 4-27 : Les noms objets du verbe *déduire* (en %)**

De manière générale, le nombre des objets nominaux est assez réduit, ce qui peut s'expliquer par la fréquence élevée de la construction en complétive *que*. Sémantiquement, les noms peuvent renvoyer soit à une conclusion obtenue à l'issue d'un processus d'inférence (*résultat*, *conséquence*), soit à un résultat non vérifié (*hypothèse*) [129] :

[129] **On peut en déduire les hypothèses suivantes**, que la comparaison des deux cohortes devrait permettre de tester. (sociologie.xml-s16747)

Dans l'exemple ci-dessus, le verbe *déduire* introduit une hypothèse, fonction qui est absente pour le verbe *conclure*.

Cette sous-classe est à rapprocher des verbes d'opinion, car si les verbes d'inférence expriment une prise de position forte de l'auteur, celle-ci a tendance d'être atténuée à l'aide de la forme modalisée *on+pouvoir*. Par ailleurs, cette sous-classe se caractérise par un effacement énonciatif (*permettre de conclure*, *il faut en déduire que*). Les deux verbes ont en commun de privilégier la construction en complétive *que* comportant souvent le pronom anaphorique *en*. Si le verbe *déduire* est exclusivement utilisé pour marquer l'issue d'une inférence, *conclure* peut traduire non seulement une inférence, mais aussi une conclusion.

Les patrons lexico-syntaxiques sont récapitulés dans le tableau 4-14 :

conclure	<p>1. [[humain=auteur, chercheur]] (pouvoir) (de [[quelque chose = source]]) conclut [complétive <i>que</i>]  Ex : <b>Cooper (1998) en a conclu que</b> les sujets en présence d'une source de fausse attribution, lors de la première étude, avaient appris à ne pas ressentir la dissonance. (psycho.xml-s10288)  Sens : Quelqu'un tire une conséquence d'un élément ou d'un ensemble d'éléments abstraits à la suite par le moyen d'un raisonnement.</p>
déduire	<p>1. [[humain=auteur, chercheur]] (pouvoir) (de [[quelque chose = source]]) déduit [complétive <i>que</i>]  Ex : <b>Peut-on</b> pour autant <b>en déduire qu'</b>une approche paysagère de la forêt est non pertinente ? (geo.xml-s10008)</p> <p>2. [[humain=auteur, chercheur]] déduit de [[quelque chose = source]] [[relation propriété thèse = résultat]]  Ex : <b>Il n'en déduit</b> pourtant pas <b>un renversement</b> de tendance, mais <b>conclue</b> à une plus grande variété des modèles d'évolution des villes. (geo.xml-s9592)</p> <p>3. [[relation propriété thèse = résultat]] (pouvoir) se déduire (être déduit) de [[quelque chose = source]]  Ex : <b>Les fonctions</b> de vraisemblance tenant compte des cas de censure <b>s'en déduisent...</b> (eco.xml-s8483)  Lexset [[relation propriété thèse = résultat]] : <i>existence, nécessité, hypothèse, conséquence, résultat</i>  Sens : Quelqu'un fait sortir d'une vérité ou d'une supposition admise comme vérité la conséquence logique qu'elle contient implicitement.</p>
Valeurs sémantiques	valeur interprétative, valeur argumentative
Routines	<i>permettre de conclure</i> (21/392) <i>conduire à conclure</i> (6/392) <i>amener à conclure</i> (5/392) <i>permettre de déduire</i> (7/176) <i>il faut conclure que</i> (11/392) <i>il faut en déduire que</i> (5/176)
Commentaires	Le pronom <i>on</i> est très fréquent.

**Tableau 4.15 : Patrons lexico-syntaxiques des verbes d'inférence**

#### 4.4.8. Récapitulatif

À l'issue de l'analyse, nous avons montré que les verbes de positionnement ont des propriétés sémantiques et syntaxiques différentes, marquant des prises de position de l'auteur fortes ou discrètes.

D'abord, les verbes d'opinion, les verbes d'inférence et les verbes d'hypothèse sont sémantiquement et syntaxiquement plus proches, servant à exprimer des prises de position marquées et fortes de l'auteur. L'étude des patrons verbaux nous permet de dégager deux caractéristiques. D'une part, ces verbes ont tendance à être modalisés. En associant la communauté de discours et le lecteur à l'argumentation, l'auteur atténue sa responsabilité et favorise la crédibilité de son assertion. Ce résultat semble

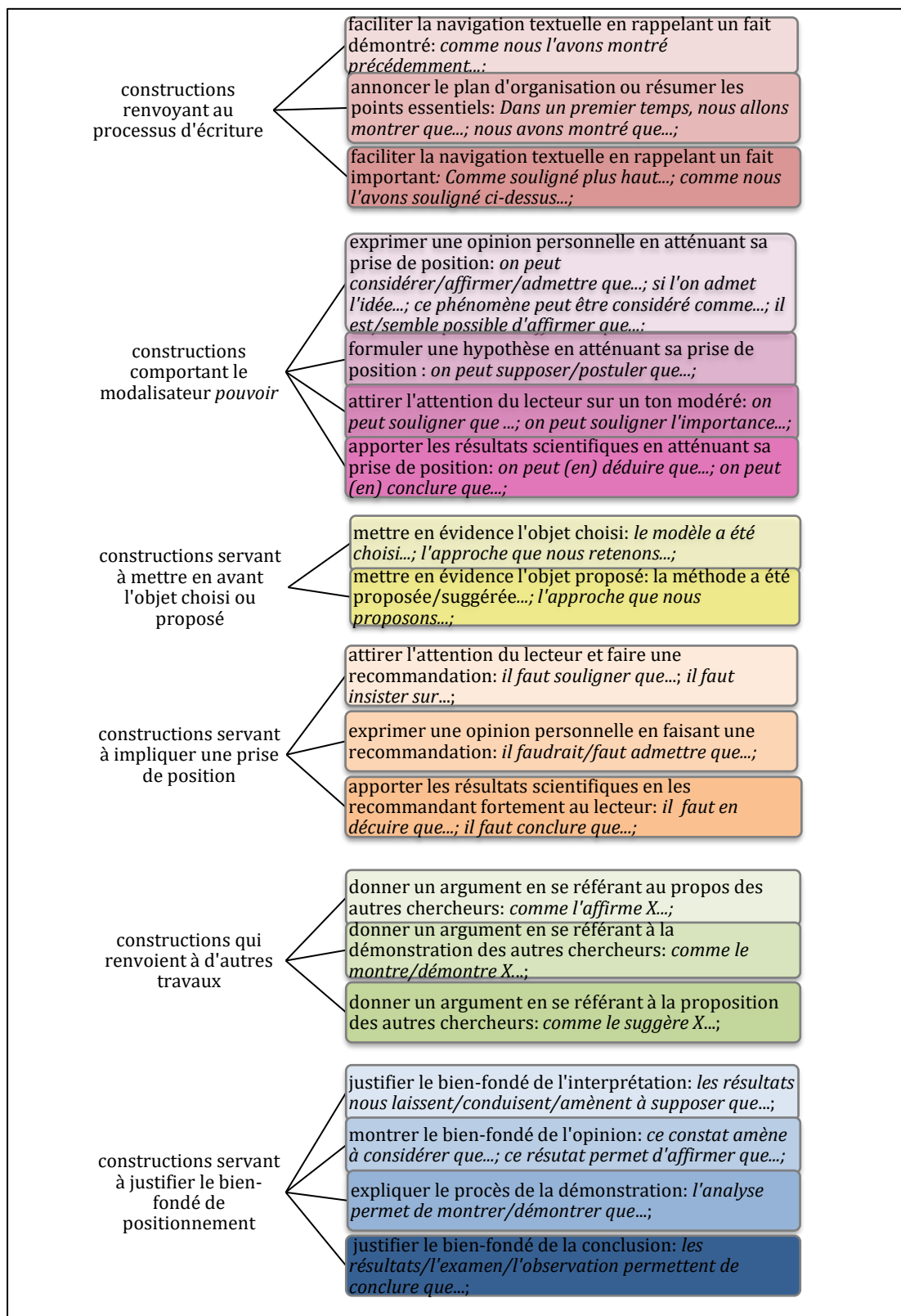
rejoindre l'étude de Tutin (2010a) selon laquelle, les verbes exprimant un positionnement fort sont souvent « tempérés par la dilution de la prise en charge de l'auteur » (*ibid.*, p. 34). D'autre part, ils ont un rôle argumentatif important, notamment lorsqu'ils sont associés aux adverbes discursifs (*cependant, en effet, aussi*) ou qu'ils apparaissent dans des constructions causatives (*permettre de considérer que, amener à supposer que*). Ces marqueurs linguistiques permettent de justifier le bien-fondé de l'assertion de l'auteur.

Ensuite, les verbes de démonstration et les verbes de mise en relief se rapprochent en ce qu'ils mettent en avant le rôle de chercheur de l'auteur (*comme nous l'avons souligné, nous montrons ici l'importance de*), contrairement aux verbes de positionnement fort qui sont fortement associés au rôle d'argumentateur (Fløttum et al., 2006). Par ailleurs, ils se caractérisent par un effacement énonciatif, à travers notamment les sujets métonymiques (*étude montre que*), les constructions impersonnelles (*il a été montré que*), les constructions avec le verbe impersonnel introducteur *falloir* (*il faut souligner que*).

Enfin, les verbes de proposition et les verbes de choix expriment un positionnement discret, ayant plus d'affinité avec la construction en complément nominal. De même, ils mettent en avant un processus d'effacement énonciatif, marqué notamment par le passif (*approche adoptée*).

De manière tendancielle, l'étude des patrons lexico-syntaxiques des verbes de positionnement a mis en relief la double contrainte dans l'expression des prises de position (qu'elles soient fortes ou discrètes) : exprimer son point de vue et viser l'objectivité. (Boch, 2013 ; Grossmann 2014a). Ainsi, l'E/A des patrons des verbes de positionnement peut permettre aux apprenants de comprendre le processus complexe des prises de position et de se former une représentation du discours scientifique. Pour ce faire, il est important de relier les patrons verbaux à leur fonction, ce qui permet d'expliquer à quoi servent ces éléments linguistiques.

Nous présentons dans la figure 4-28 les constructions verbales identifiées de notre analyse, celles-ci sont organisées selon leur fonction rhétorique.



**Figure 4-28 : Les principales constructions verbales des verbes de positionnement et leurs fonctions**

À l'issue de l'analyse des patrons de deux classes de verbes, nous pensons avoir

fait émerger certaines spécificités liées au genre de l'écrit scientifique que nous allons mettre en avant dans le bilan ci-dessous.

## 4.5. Bilan

Dans ce chapitre, nous avons étudié les verbes d'analyse et de positionnement à travers leurs constructions syntaxiques associées aux valeurs sémantiques, les pronoms sujets mettant en jeu l'auteur ainsi que les arguments nominaux. L'ensemble de ces informations a été intégré dans les patrons lexico-syntaxiques. Ce faisant, la modélisation des patrons nous a permis de mettre en évidence les propriétés sémantiques et syntaxiques liées à ces deux classes de verbes. Si les verbes d'analyse mettent en place l'objectivité du processus de recherche à travers notamment la thématization de l'objet des études et la justification du choix, les verbes de positionnement mobilisent à la fois les moyens de l'argumentation et de l'effacement énonciatif, constituant ainsi un processus paradoxal (Boch, 2013). Les spécificités qui émergent à l'issue de l'analyse des verbes peuvent correspondre à trois dimensions :

- dimension discursive : On entend ici par ce terme la relation entre l'auteur et le lecteur et celle entre l'auteur et son texte (Hyland, 2000). Par exemple, la construction en *comme* utilisée pour faciliter la navigation textuelle et l'emploi inclusif du pronom *on*.
- dimension énonciative : La mention de l'auteur est souvent implicite si l'on considère les cas du passif, des différents types de constructions impersonnelles et les emplois métonymiques. Cependant, l'auteur est souvent visible dans les rôles du scripteur et du chercheur. Par exemple, lorsque les verbes revêtent une valeur métatextuelle.
- dimension argumentative : L'argumentation consiste à persuader son lectorat (Hyland, 2005a), ceci se manifeste notamment à travers le référencement et le raisonnement pendant notre analyse. Par exemple, le renvoi à d'autres opinions où les verbes prennent comme sujet un nom représentant d'autres chercheurs, la présence non négligeable des constructions causatives et des adverbes discursifs exprimant une concession (*cependant*), une opposition (*par contre*), un ajout (*par ailleurs*), une reconsidération (*en effet*), etc. Ainsi, comme le souligne Pollet

(2014), dans l'écrit scientifique, l'auteur étaye et prouve ses affirmations à partir de différentes preuves et de différentes voix énonciatives pour construire son raisonnement argumenté.

Enfin, à l'issue de l'analyse, la question que nous nous posons est de savoir comment aborder les patrons lexico-syntaxiques des verbes véhiculant les spécificités liées au genre de l'écrit scientifique dans la classe de FLE.

Avant de réfléchir à cette question, il nous semble utile d'examiner les difficultés des apprenants en matière de verbes étudiés dans le chapitre V afin de proposer des pistes didactiques adéquates.

---

## Chapitre V.

# Diagnostic des besoins langagiers chez les apprenants chinois

---

5.1. TROIS NOTIONS ESSENTIELLES POUR LE DIAGNOSTIC : LES NORMES, LES ERREURS ET LES SPECIFICITES.....	213
5.1.1. Les normes linguistiques propres aux écrits scientifiques.....	213
5.1.2. Les erreurs.....	214
5.1.3. Les spécificités .....	215
5.2. DIAGNOSTIC I : ERREURS RELATIVES AUX VERBES DU LST .....	216
5.2.1. Typologie des erreurs en FLM et FLE.....	216
5.2.2. Constitution d'une typologie des erreurs centrée sur les verbes du LST .....	222
5.2.3. Analyse des erreurs.....	225
5.3. DIAGNOSTIC II : SPECIFICITES LIEES AUX CONSTRUCTIONS VERBALES DU LST .....	240
5.3.1. Études représentatives sur les spécificités dans le domaine de l'EAP .....	240
5.3.2. Constitution de la grille de diagnostic sur les spécificités .....	247
5.3.3. Analyse des spécificités .....	250
5.4. RECAPITULATIF DU DIAGNOSTIC .....	271
5.5. PISTES DIDACTIQUES .....	273
5.5.1. Enseignement des patrons avec le corpus .....	274
5.5.2. Deux expérimentations initiales .....	277
5.5.3. Exemples d'exercices adaptés.....	286

---

Dans le chapitre précédent, nous avons étudié les verbes d'analyse et de positionnement en proposant une approche inspirée du modèle CPA de Hanks. Les résultats de l'analyse, modélisés dans une base de données, nous permettront de proposer des activités pour l'E/A des patrons verbaux dans deux expérimentations didactiques présentées à la fin de ce chapitre. Cependant, avant de mener une réflexion didactique, il nous semble important d'établir un diagnostic, une analyse des besoins langagiers liés à l'utilisation de ces verbes chez les apprenants afin de mieux cibler leurs difficultés et de proposer par la suite des solutions appropriées aux besoins.

Ce diagnostic comporte deux parties. La première vise à révéler les erreurs liées aux constructions verbales dans un corpus d'apprenants. La deuxième partie, fondée sur la comparaison entre un corpus d'apprenants et un corpus d'étudiants natifs, porte sur les spécificités et consiste à mettre en évidence les cas de suremploi et de sous-emploi de verbes du LST et de constructions verbales (en particulier, les verbes d'analyse et de positionnement) dans notre corpus d'apprenants chinois. Notre analyse ne se limite pas aux aspects lexicaux et syntaxiques, mais intègre aussi le niveau

discursif. Les questions sous-jacentes auxquelles nous essaierons de répondre sont les suivantes : les apprenants chinois maîtrisent-ils les verbes du LST ? Plus précisément, concernant les verbes étudiés, maîtrisent-ils la diversité des constructions verbales ? Quelles sont leurs lacunes liées aux constructions verbales ? Enfin, au niveau discursif, emploient-ils les verbes pour décrire leur objet de recherche, pour se référer aux travaux d'autrui, pour expliquer et démontrer les faits, pour exprimer leur prise de position ? Dans l'affirmative, s'ils ont des lacunes, quelles sont-elles ?

Ce chapitre sera organisé comme suit. Nous aborderons d'abord les notions d'erreur, de spécificité et de norme que nous définissons dans le cadre de cette étude. Une fois ces notions précisées, nous présenterons la première partie de notre diagnostic portant sur les erreurs, qui sera guidé par une typologie d'erreurs, en partie fondée sur un état de l'art. Nous effectuerons une analyse des erreurs des apprenants chinois en nous focalisant sur les aspects lexicaux, sémantiques, syntaxiques et discursifs qui permettent d'évaluer le degré de maîtrise ou non des constructions verbales dans l'écrit scientifique par les apprenants. Dans un second temps, nous proposerons un diagnostic des spécificités, en mettant en contraste deux corpus (le Corpus Sup-Natifs et le Corpus Sup-Chinois) avec l'outil TXM. Nous confronterons également les écrits des apprenants chinois à nos patrons verbaux afin d'en dégager les spécificités. Enfin, après avoir proposé un bilan sur les besoins langagiers, nous présenterons deux expérimentations didactiques à partir des résultats d'analyse en menant une réflexion sur l'E/A des patrons verbaux avec corpus.



## 5.1. Trois notions essentielles pour le diagnostic : les normes, les erreurs et les spécificités

Avant de commencer le diagnostic des difficultés des apprenants, il importe de revenir sur quelques notions importantes telles que norme, erreur et spécificité, sur lesquelles s'appuie notre étude. Parmi ces notions, nous nous attachons plus particulièrement à la notion de spécificité qui est au cœur de notre étude.

### 5.1.1. Les normes linguistiques propres aux écrits scientifiques

Plusieurs études (Anctil, 2010 ; Ying, 2013) s'accordent sur la complexité à définir des normes linguistiques. Anctil (*ibid.*) a dressé dans sa thèse un aperçu général des différents types de normes que nous pouvons dégager selon trois critères essentiels : 1) l'usage 2) les règles grammaticales 3) le contexte de communication et le genre de discours. Dans le cadre de cette thèse, nous ferons appel à ces trois critères pour définir les normes linguistiques.

D'abord, les normes peuvent être déterminées par l'usage. Il s'agit de rendre compte des régularités observables de la langue à travers la fréquence et des tendances statistiques. Ensuite, les normes peuvent être imposées par des institutions comme l'Académie française à travers des « ouvrages de référence comme les grammaires, mais surtout à travers les dictionnaires » (*ibid.*, p. 55). Enfin, les normes varient en fonction des situations de communication et des genres de discours. Ainsi, le contexte académique exigera l'emploi d'un vocabulaire formel ou spécialisé, en conformité avec les pratiques langagières dans le milieu de l'E/A et de la discipline universitaire.

Dans le contexte des écrits scientifiques, les normes linguistiques sont conçues dans la communauté de discours où elles sont appliquées. Comme le souligne Cavalla (2010a), elles renvoient aux codes linguistiques reconnus par tous les membres de la communauté en question. Ces codes décrits en termes de discours (l'effacement du sujet, l'utilisation des tournures nominales et impersonnelles) (Reuter, 1998) et de lexique (collocation, phraséologie transversale) (Chavez, 2008 ; Cavalla, 2015) peuvent « engendrer des déconvenues rédactionnelles que le lecteur repère

rapidement et tend à expertiser négativement quand il est spécialiste » (Cavalla, 2010a, p. 208). En ce qui concerne notre étude, les patrons verbaux relevés dans le corpus d'experts Scientext-SHS présentent des régularités et l'usage courant du genre de l'écrit scientifique (*cf.* Chapitre IV). Ils pourraient donc constituer les formes « normales » de l'usage de l'écrit scientifique. Cette vision de norme reliée au genre de discours nous permet également de considérer les écrits universitaires des étudiants natifs comme s'écartant de la norme en ce qu'ils ne maîtrisent pas complètement les codes de l'écriture scientifique (Rinck et al., 2007 ; Rinck, 2010b ; Boch & Rinck, 2016).

Ainsi, il nous semble pertinent de choisir les experts comme nos locuteurs de référence. La conception des normes linguistiques que nous avons adoptée se traduit par le choix du corpus d'experts Scientext-SHS comme modèle.

### 5.1.2. Les erreurs

L'erreur est souvent définie comme « un écart par rapport à une norme » (Anctil, 2005, p. 66). Ainsi la détermination de l'erreur varie en fonction des normes choisies. Il apparaît important de comprendre ce qu'est l'erreur dans le cadre de cette thèse. Tenant compte de la conception des normes linguistiques présentées ci-dessus, nous pouvons définir l'erreur linguistique propre aux écrits scientifiques comme suit :

<i>Élément linguistique (mot, expression, ou construction) qui ne respecte pas soit les règles grammaticales de la langue soit les codes linguistiques reconnus par la communauté de discours.</i>
--

Il faut souligner que la définition adoptée ici se limite à la production des discours scientifiques. Elle nous permet de désigner tant les problèmes au niveau linguistique (morphologique, sémantique, syntaxique) que les emplois marqués par rapport au genre de discours, qui pourraient être considérés plutôt comme des maladroites.

### 5.1.3. Les spécificités

Dans le chapitre II, nous avons expliqué que l'interlangue se caractérise non seulement par les erreurs mais aussi par des cas de suremploi et de sous-emploi (Granger, 2007). Nous pourrions faire appel au terme « spécificités » pour désigner ces cas de suremploi et de sous-emploi qui « associés aux erreurs pures et simples, trahissent l'origine allophone des apprenants » (*ibid.*, p. 132).

On peut souligner que, contrairement aux erreurs qui sont souvent définies par rapport aux règles grammaticales, les spécificités sont plus subjectives et dépendent du jugement sur l'usage, sur le contexte de communication et sur le genre de discours. Concernant notre étude, il s'agit des spécificités se rapportant aux codes linguistiques, plus précisément aux patrons verbaux (*cf.* Chapitre IV), propres au genre de l'écrit scientifique. En d'autres termes, nous entendons par spécificités :

*Les cas de sous-emploi ou même d'absence des patrons verbaux (codes spécifiques au genre des écrits scientifiques attendus par les locuteurs natifs spécialistes du domaine) et les cas de suremploi de mots, d'expressions et de structures non reconnus comme faisant partie de la communauté de discours.*

Rappelons que l'interprétation de spécificités doit reposer sur la comparaison entre le corpus d'apprenants et un corpus de locuteurs natifs de type plus ou moins équivalent (*cf.* Chapitre II).

En construisant leur interlangue, les apprenants peuvent commettre des erreurs grammaticales, faire des traductions spécifiques et sous-employer des mots ou des expressions dans leur texte. Ces cas se cumulent et se rencontrent fréquemment dans un même paragraphe. En ce sens, de notre point de vue, seule l'analyse intégrant les erreurs et les spécificités permet de distinguer les écrits des apprenants de ceux des locuteurs natifs et de faire apparaître les enseignements à mettre en place auprès des apprenants pour leurs écrits académiques.

Dans la section suivante, nous allons commencer par les analyses des erreurs. Les analyses des spécificités seront abordées dans la section 5.3.

## 5.2. Diagnostic I : erreurs relatives aux verbes du LST

Dans cette partie, nous cherchons à établir le diagnostic sur les erreurs. Pour ce faire, nous dresserons d'abord un état de l'art concernant les analyses des erreurs afin de proposer une typologie des erreurs. Elle nous aidera ensuite à relever les erreurs fréquentes dans les écrits des apprenants chinois aux plans morphologique, sémantique et syntaxique.

### 5.2.1. Typologie des erreurs en FLM et FLE

Dans le chapitre II, nous avons montré que les corpus d'apprenants sont souvent utilisés pour l'analyse des erreurs assistée ou non par ordinateur. Ces études s'avèrent utiles pour comprendre et analyser les productions langagières des apprenants en raison de leur taille et de leur représentativité (Granger, 2012 ; De Cock & Tyne, 2014).

En fait, comme le souligne Luste-Chaâ (2008), l'analyse des erreurs a pris son essor dans les années 70 et différentes typologies fondées sur des critères de classement très divers sont proposées, notamment en contexte de productions d'écrits (Catach *et al.*, 1980 ; Garcia-Debanc, 1990 ; Granger & Monfort, 1994 ; Anctil, 2010). Cette évolution s'est accompagnée de l'apparition des corpus d'apprenants et de la linguistique de corpus, notamment pour le traitement des données (par exemple les corpus pourvus d'annotations d'erreurs).

Selon Besse et Porquier (1991), l'analyse des erreurs à l'aide de corpus d'apprenants a essentiellement un double objectif : 1) au niveau théorique, elle permet de mieux comprendre les processus d'apprentissage d'une langue étrangère ; 2) du point de vue didactique, elle permet de comprendre les difficultés des apprenants et d'améliorer l'enseignement. Notre étude s'inscrit donc dans cette deuxième dimension diagnostique/didactique.

Étant donné que notre intérêt porte sur le lexique, nous présenterons d'abord la typologie des erreurs lexicales d'Anctil (2005, 2010) construite dans une perspective didactique avec un public du français langue maternelle (FLM). Dans un deuxième

temps, en vue de mieux cibler notre public, nous proposons un aperçu des erreurs fréquentes chez les apprenants chinois, notamment en matière de lexique et de syntaxe.

### 5.2.1.1. Typologie d'Anctil

La typologie d'Anctil (2005) nous intéresse particulièrement par son entrée lexicale et son analyse détaillée et systématique. Elle est établie à partir du recensement des erreurs lexicales dans un corpus de 103 textes narratifs/explicatifs produits par les élèves francophones de troisième secondaire au Québec (14-15 ans), d'un niveau équivalent aux élèves de troisième année de collège en France. Ici, le terme 'erreurs lexicales' est défini dans un sens large, qui englobe toutes les erreurs relatives au lexique, d'ordre orthographique, sémantique et sur les propriétés combinatoires<sup>1</sup>. Le lexique, considéré comme élément organisateur de la langue est étroitement associé à d'autres domaines comme la grammaire, la syntaxe et la sémantique. À l'instar d'Anctil, nous pensons que le lexique ne peut pas être étudié isolément mais doit être présenté comme lié à d'autres compétences linguistiques.

Cette typologie est divisée en trois catégories : 1) les aspects liés à l'unité lexicale (forme, sens, propriétés de combinatoire grammaticale/lexicale), 2) l'utilisation d'une lexie fictive (création d'un sens inexistant, impropriété, dérivation morphologique inexistante) et 3) les erreurs liées au contexte (registre de langue, répétition, imprécision, etc.) (Anctil, 2005, p. 65-66). Néanmoins, on peut remettre en question certains regroupements. Par exemple, des erreurs sémantiques sont réparties dans plusieurs cases de la grille (sens, création d'un sens inexistant, impropriété, imprécision). Ces problèmes ont été ensuite pris en compte par l'auteur et des aménagements ont été apportés dans la nouvelle version. Dans cette dernière, Anctil (2010) rassemble les erreurs en trois grandes classes telles que forme, sens, et combinatoire restreinte aux niveaux grammatical, lexical et pragmatique.

---

<sup>1</sup> Concernant les propriétés de combinatoire, Anctil (2005) a suivi les principes de la lexicologie de Mel'čuk (Lexicologie Explicative et Combinatoire) et a distingué deux grands axes : la combinatoire grammaticale et la combinatoire lexicale. Le premier relève du comportement morphologique et syntaxique d'une unité lexicale (lexème ou locution) alors que le deuxième renvoie aux problèmes de collocation.

La nouvelle propose une classification plus homogène des erreurs sémantiques. Elle couvre quasiment la totalité des types d'erreurs lexicales et s'avère intéressante pour l'analyse des erreurs liées aux aspects sémantique et combinatoire. C'est pourquoi, bien qu'elle soit élaborée avec un public de FLM, nous pensons qu'elle peut être opérationnelle pour le FLE.

Un point intéressant dans l'analyse effectuée par l'auteur, est qu'il est montré qu'à travers la répartition des classes de mots dans chaque grande classe d'erreurs, le verbe représente 55,5% des erreurs de combinatoire syntaxique, 41,5% des erreurs de sens et 30,7% des erreurs formelles (Anctil, 2010, p. 280). Nous pensons qu'il est indispensable de prendre en compte ces classes lors de l'analyse des erreurs liées au verbe.

Nous retenons de cette typologie qu'il est important de prendre en compte les aspects de combinatoire restreinte (aux niveaux grammatical, lexical et pragmatique). Les différents types d'erreurs sont illustrés dans le tableau 5.1 :

<b>COMBINATOIRE RESTREINTE</b>	<b>GRAMMATICALE</b>	<b>SYNTAXIQUE (RÉGIME)</b>	Absence d'un complément obligatoire de l'UL
			Emploi transitif indirect d'un verbe transitif direct
			Utilisation d'une préposition/conjonction erronée
			Préposition régie absente
			Emploi transitif d'un verbe intransitif
			Construction erronée dans un complément
		<b>MORPHO-SYNTAXIQUE</b>	Non-respect de la classe de mots de l'UL
			Non-respect du genre du nom
			Non-respect du caractère invariable de l'UL
			Utilisation d'un verbe auxiliaire erroné
			Classe résiduelle
	<b>LEXICALE</b>	Choix d'un collocatif inapproprié (collocation inexistante)	
		Choix d'un collocatif valable, mais mauvaise fonction lexicale	
		Non-respect des contraintes d'emploi du collocatif	
		Erreur de dérivation sémantique	
	<b>PRAGMATIQUE</b>	Emploi d'une UL appartenant à un registre de langue inapproprié	
Emploi d'une UL véhiculant un contenu connotationnel incompatible avec le contexte			
Emploi d'une UL appartenant à une variété dialectale incompatible avec le contexte			

**Tableau 5.1 : Les types d'erreurs de combinatoire restreinte (Anctil, 2010, p. 194)**

Comme on peut le constater dans le tableau 5.1, en ce qui concerne la syntaxe, plusieurs sous-classes apparaissent intéressantes, par exemple, « préposition régie absente », « construction erronée dans un complément ». En revanche, les erreurs liées à l'ordre des mots ne sont pas intégrées, ce qui s'explique par le fait que l'auteur travaille sur des productions de natifs qui commettent rarement ce type d'erreurs. Nous allons intégrer ces erreurs dans notre typologie puisque l'ordre des mots constitue un

aspect important pour le diagnostic sur les constructions syntaxiques.

Par ailleurs, les erreurs sur les propriétés de combinatoire lexicale sont divisées en trois sous-classes liées aux collocations. Il nous semble important de les intégrer dans notre typologie parce que les collocations constituent une partie importante de patrons verbaux. Cependant, nous allons repenser ces sous-classes en fonction de notre objet d'analyse pour voir si elles sont adaptées aux apprenants chinois.

Dans la section suivante, nous présentons deux thèses consacrées à l'analyse des erreurs fréquentes chez les apprenants chinois.

### **5.2.1.2. Études sur les erreurs des apprenants chinois en FLE**

Dans le domaine du FLE, les erreurs commises par les apprenants chinois sont peu étudiées. Nous pouvons par contre mentionner deux thèses qui proposent des analyses sur les principales erreurs commises à l'écrit par les apprenants chinois. Il est intéressant de noter que les deux auteurs ont adopté deux approches différentes : la thèse de Hou (2008) adopte une approche comparative et vise à comprendre les sources des erreurs tandis que la thèse de Ying (2013), avec une approche cognitive, cherche à proposer une remédiation au plan didactique.

Hou (2008) a recensé 4498 erreurs dans les textes (à dominante argumentative) collectés auprès d'étudiants de français dans deux universités chinoises dont le niveau de langue n'est pas communiqué. Son étude se fonde sur une approche comparative. L'objectif est de chercher la source des erreurs à travers la comparaison entre la langue maternelle (le chinois) et la langue étrangère (le français).

Les erreurs relevées sont classées dans trois grandes catégories : erreurs de forme (orthographe, conjugaison, etc.), d'ordre syntactico-sémantique (préposition, structure, synonymie, etc.) et d'ordre pragmatique (registre, coréférence, etc.). Pour chaque type d'erreurs catégorisé, l'auteur a proposé une analyse qui met en avant le fait que les erreurs liées à la langue maternelle ne constituent pas la seule cause d'erreurs chez les apprenants, et que les caractéristiques du français ainsi que les stratégies d'apprentissage des apprenants jouent un rôle important.

D'après les statistiques obtenues, les erreurs les plus fréquentes commises par les

apprenants chinois sont liées à la forme (orthographe et genre) (34,9%) et à l'utilisation de l'article (16,8%). Concernant les constructions verbales, l'auteur a notamment souligné deux erreurs :

- (1) les erreurs de prépositions, par exemple, les apprenants n'utilisent pas correctement la préposition régie par le verbe (ex : *réfléchir à, continuer de*). D'après l'auteur, ces erreurs sont dues essentiellement à la propriété du verbe (le caractère transitif ou intransitif) ;
- (2) les erreurs de structure (ordre des mots, emploi de la négation, etc.), par exemple, les apprenants mettent souvent l'adverbe avant le verbe dans la phrase, du fait d'un transfert de la langue maternelle. En fait, comme le souligne l'auteur, les erreurs de structure sont souvent liées aux caractéristiques du français, mais également à celles du chinois. Autrement dit, l'influence de la langue maternelle induit des erreurs chez les apprenants.

Nous pensons que les erreurs de préposition et de structure sont deux aspects importants à intégrer dans notre analyse.

Dans une autre étude, Ying (2013) a constitué un corpus de 1020 erreurs auprès de deux publics : des apprenants à l'Alliance Française de Wuhan (niveau débutant) et des étudiants de français à l'Université Jiangnan (niveau intermédiaire). Ses données sont des exercices de composition écrite (à dominante descriptive et argumentative). Trois grandes catégories ont été proposées : erreurs lexicales, grammaticales et sociopragmatiques, comme l'illustre le tableau 5.2 :

Type lexical	Type grammatical	Type sociopragmatique
Manifestation : erreur de mots	Manifestation : violation d'une règle grammaticale	Manifestation : violation d'une règle de conduite sociale ou d'une règle d'usage de la langue
Deux sous-types : - erreur de forme - erreur de concept	Trois sous-types : - Erreur morphologique dérivationnelle et flexionnelle - syntaxique - Sémantique	

**Tableau 5.2 : Classement des erreurs proposé par Ying (2013, p. 94)**

Concernant les erreurs fréquentes relevées, les résultats obtenus semblent correspondre partiellement à ceux de Hou. D'une part, comme dans l'étude de Hou,



les erreurs de forme (orthographe, genre) (16,86%) et d'article (9,8%) apparaissent fréquentes. D'autre part, l'auteure a mis en avant un nombre non négligeable des erreurs de TAM (temps, aspect et modalité) du verbe (13,33%) et de sens (11,47%). Par exemple, concernant le sens, il a été souligné que les apprenants chinois confondent les verbes *distribuer* et *affecter*, et ne les utilisent pas dans le contexte approprié. Cette confusion sémantique concerne de nombreux verbes synonymes ou ayant des sens proches ou complémentaires tels que *trouver/chercher*, *amener/apporter*, *participer/assister/adhérer*. En effet, la plupart de ces verbes ne sont interchangeable que dans certains contextes précis. Étant donné que les apprenants chinois apprennent souvent ces verbes hors contexte, sans prendre en compte leurs associations lexicales ou leurs préférences syntaxiques, ils commettent souvent des erreurs de sens, d'où l'importance d'aborder les propriétés de combinatoire dans l'enseignement.

En dehors des erreurs syntaxiques similaires à celles relevées par Hou (en ce qui concerne les prépositions régies par le verbe et l'ordre des mots), Ying a également classé les erreurs liées aux constructions erronées telles que les phrases incomplètes ou mal composées. Par exemple, l'omission des compléments d'objet comme avec le verbe *renoncer* (*à ce rêve*), ou la formulation de phrases incompréhensibles produites par une traduction littérale du chinois vers le français. Ces aspects concernant les constructions verbales seront pris en compte dans notre grille d'analyse.

Par ailleurs, elle a suggéré quelques remédiations vis à vis des erreurs syntaxiques. Par exemple, concernant l'apprentissage du verbe, l'auteure (*ibid.*, p. 226) a souligné que celui-ci doit inclure une connaissance sur la propriété et l'utilisation du verbe, proposant une mémorisation des verbes en compagnie des prépositions. Son point de vue rejoint celui de Chollet et Robert (2007) qui ont proposé un ouvrage répertoriant les verbes français et leurs prépositions pour l'E/A du FLE. Dans la même perspective, nous pensons que l'E/A des verbes en compagnie des prépositions correspond parfaitement à notre conception de patron dans laquelle un mot doit s'apprendre en compagnie des éléments linguistiques fréquents qui l'entourent. Notons par ailleurs que ce point de vue est partagé par de nombreux travaux sur la didactique du lexique dans lesquels nous pouvons mentionner l'approche lexicale (The lexical approach) proposée par Lewis (1993 ; 1997).

Enfin, il apparaît que les résultats de Hou et de Ying ont notamment mis en évidence des maladroites liées aux connaissances élémentaires du français telles que l'orthographe, la conjugaison, l'accord, etc. Les erreurs au niveau des constructions verbales se montrent en revanche beaucoup moins fréquentes. Ces résultats pourraient être liés aux niveaux des apprenants et aux types des textes du corpus. Mais ceci ne veut pas pour autant dire que les constructions verbales posent moins de problèmes aux apprenants chinois. Bien au contraire, on peut supposer qu'à un niveau plus élevé (à partir du niveau B2), les apprenants feront moins d'erreurs sur la forme et que davantage de difficultés apparaîtront aux niveaux phrastique et discursif. Par ailleurs, les outils de correction orthographique et grammaticale peuvent relever la plupart des erreurs liées aux règles grammaticales mais difficilement celles qui apparaissent aux niveaux lexico-syntaxiques et discursifs. On suppose donc que les erreurs de constructions verbales ne peuvent pas être détectées par la machine et constituent alors une des grandes difficultés des apprenants chinois.

Voyons désormais ce que nous retenons pour la constitution de notre typologie des erreurs.

### **5.2.2. Constitution d'une typologie des erreurs centrée sur les verbes du LST**

Après avoir analysé les travaux de quelques auteurs à propos de l'analyse des erreurs dans le domaine du FLE et du FLM, nous avons retenu plusieurs points intéressants à prendre en compte dans la construction de notre typologie des erreurs :

- (1) les différents aspects liés au verbe : forme, sens, propriétés combinatoires
- (2) les erreurs fréquentes liées à l'emploi du verbe chez les apprenants chinois : erreurs de TAM, erreurs de sens, erreurs de préposition régie par le verbe
- (3) concernant l'emploi des constructions verbales chez les apprenants chinois : ordre des mots, emploi de négation, constructions incomplètes ou erronées

Par ailleurs, nous avons effectué les premières observations dans le corpus. Nous avons constaté que certaines erreurs sont quasiment absentes de notre corpus. Par

exemple, les erreurs d'emploi de négation liées aux verbes du LST (ex : \* *Aucun candidat a été retenu*) ne sont pas étudiées car absentes du corpus.

Notons par ailleurs que notre typologie joue un rôle de guidage dans l'analyse et que des modifications pourraient être apportées en fonction des résultats de diagnostic.

Rappelons que notre objectif est de constituer une grille qui nous aidera à diagnostiquer les difficultés des apprenants chinois dans l'utilisation des verbes du LST et des constructions verbales. C'est pourquoi notre typologie doit être centrée sur les verbes du LST et leurs constructions verbales, et présenter à la fois une certaine clarté et une simplicité pour une utilisation par des non spécialistes (apprenants, enseignants non linguistes).

En ce qui concerne la structure, nous distinguons dans un premier niveau quatre classes d'erreurs en fonction de catégories linguistiques : morphologique, sémantique, syntaxique et discursif, qui se subdivisent dans un deuxième niveau. Enfin, des exemples tirés du corpus sont donnés pour illustrer chaque type d'erreurs (voir le tableau 5.3).

Classes d'erreurs	Types d'erreurs	Exemples
Au niveau morphologique	Erreurs de conjugaison	Ex : Si on parle d'un sportif noir, jamais on le <b>présenter</b> (présente) comme... (socioculturel 4)
	Erreurs d'orthographe	Ex : Les poèmes <b>mentionés</b> (mentionnés) sont empruntés au...(littérature 1)
	Erreurs de temps/d'aspect/de mode des verbes (TAM)	Ex : Ensuite, l'auteur <b>ira combiner</b> (va combiner) les deux théories de traduction...
Au niveau sémantique et des co-occurrences lexicales	Erreurs de sens (sens fictif, mauvais synonyme)	Ex : On peut <b>constater</b> (voir) certains documents sur...(traduction 1)
	Erreurs de collocation (noms inappropriés ou génériques)	Ex : La plupart des critiques peuvent <b>décrire les dissimilitudes</b> (décrire les différences) entre les deux versions...(traduction 2)
	Erreurs de co-occurrence (adverbes)	Ex : Nous <b>proposerons grossièrement</b> (sommairement) une distinction de quatre grands courants : textuel, anaphorique, inactuel et aspectuel. (linguistique_1)
Au niveau de la grammaire et de la syntaxe	Erreurs d'ordre des mots (pronom, adverbe)	Ex : Tout au long du roman, l'auteur <b>a nous décrit</b> (nous a décrit) neuf voyages...
	Utilisation erronée d'une préposition	Ex : Nous l' <b>expliquons par</b> (dans) la paragraphe suivante...
	Structure syntaxique erronée	Ex : On pourrait <b>constater</b> (constater qu') il

	(omission, structure inexistante, maladroite)	y a les liens étroits entre...(littérature_2)
Au niveau du registre	Expression et structure appartenant à un registre de langue inapproprié (usage oral)	Ex : <b>ça indique</b> que le traducteur surestime le niveau (ceci indique)...(traduction_9)

**Tableau 5.3 : Typologie des erreurs centrée sur les verbes du LST**

Il faut noter que, parmi les quatre classes d'erreurs présentées dans le tableau, nous nous focalisons particulièrement sur les trois dernières. En effet, la première observation du Corpus Sup-Chinois nous révèle que les erreurs d'orthographe et de conjugaison liées au verbe sont peu présentes, parce que d'une part, notre corpus est partiellement corrigé (*cf.* Chapitre II), d'autre part, on peut supposer qu'à un niveau élevé (B1+), les erreurs au niveau morphologique diminuent significativement.

Quant aux erreurs relatives aux structures syntaxiques erronées, nous adoptons une conception plus large qui vise à prendre en compte les choix lexicaux inappropriés aux écrits scientifiques, par exemple les constructions ne respectant pas les codes linguistiques de l'écrit scientifique (ex : *Tout cela vaut la peine d'être étudié ; il est donc conseillé de l'expliquer*). Selon Anctil (2005, p. 118), ces erreurs soumises à une norme « stylistique » sont plus subjectives et de ce fait, elles sont souvent considérées plutôt comme des maladroites. Cependant, ces dernières ont reçu le statut d'erreur dans notre étude. En effet, dans la perspective de l'E/A des écrits universitaires, d'une part, ces erreurs de norme indiquent une certaine méconnaissance du genre scientifique pour les étudiants chinois ; d'autre part, elles sont souvent pénalisées par les enseignants lors des corrections. Nous illustrons plus loin ces erreurs avec des exemples plus concrets.

Enfin, il nous semble important de souligner que les types d'erreurs proposés ne sont pas assez fins. Avec un corpus de petite taille, nous préférons travailler avec des classes d'erreurs englobantes afin d'observer les principales difficultés des apprenants. D'ailleurs, nous n'avons pas intégré les erreurs comme celles de ponctuation dans notre typologie, mais nous pensons que ces erreurs sont importantes à prendre en compte dans un futur travail.

Dans la section suivante, en nous appuyant sur cette typologie, nous allons analyser les différents types d'erreurs dans le Corpus Sup-Chinois avec l'outil TXM, et tant aux niveaux quantitatif que qualitatif.

### 5.2.3. Analyse des erreurs

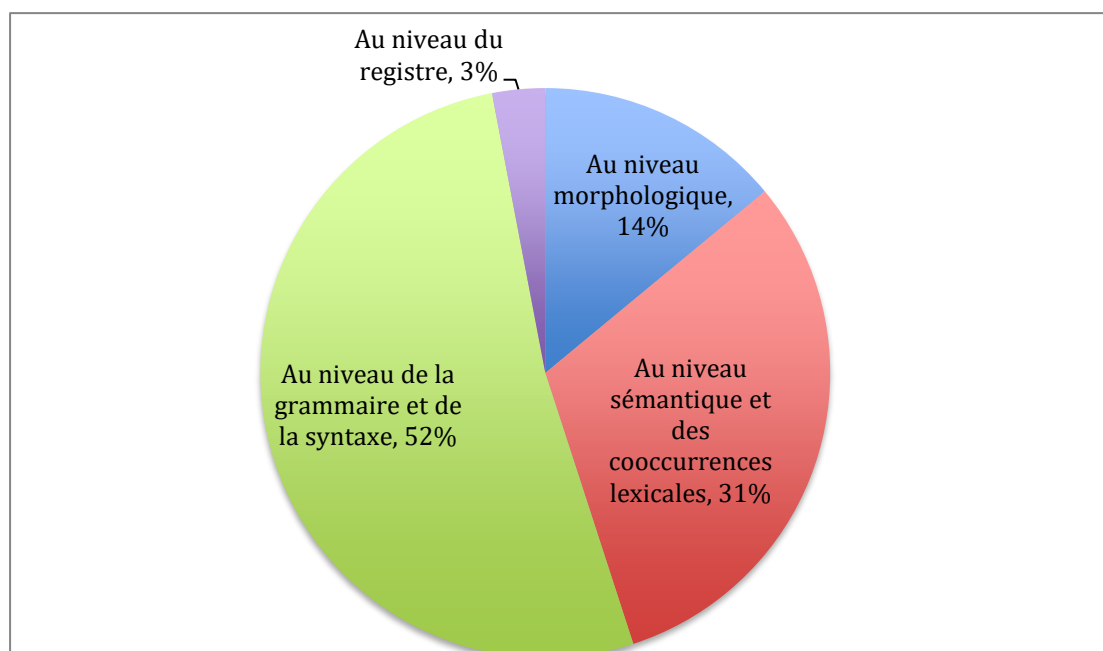
Nous nous attachons à examiner les erreurs centrées sur les verbes du LST. Pour ce faire, nous prenons 10 verbes que nous avons déjà étudiés au chapitre IV (*choisir, considérer, comparer, décrire, étudier, expliquer, montrer, noter, proposer, supposer*). Ces verbes ont été retenus parce qu'ils sont fréquents et présentent des constructions/fonctions importantes et variées dans l'écrit scientifique (*cf.* Chapitre IV). Notre objectif est donc de voir si ces verbes et leurs constructions verbales sont maîtrisés par les apprenants chinois. Si cela n'est pas le cas, quelles sont les lacunes langagières ?

Nous avons extrait les concordances à l'aide du logiciel TXM et observé manuellement les erreurs. Nous avons relevé au total 221 erreurs pour les 10 verbes retenus. Il est frappant de constater que les erreurs sont nombreuses dans les mémoires partiellement corrigés. Il y a peut-être lieu de penser que certains enseignants sont tolérants quant aux erreurs de langue et proposent alors peu de corrections. Notons par ailleurs que les erreurs d'orthographe ont été relevées avec des expressions régulières<sup>2</sup>.

La répartition des erreurs par classe est illustrée dans la figure 5-1 :

---

<sup>2</sup> Le logiciel nous propose trois expressions régulières. Le point « . » peut remplacer n'importe quel caractère ; l'étoile « \* » s'applique à ce qui précède et veut dire "de 0 à infini" ; le plus « + » s'applique à ce qui précède et veut dire "de 1 à infini". Par exemple, la requête « [lemma="c.\*er"] » permet de relever tous les mots qui commencent par la lettre « c » et qui se terminent par « er ».

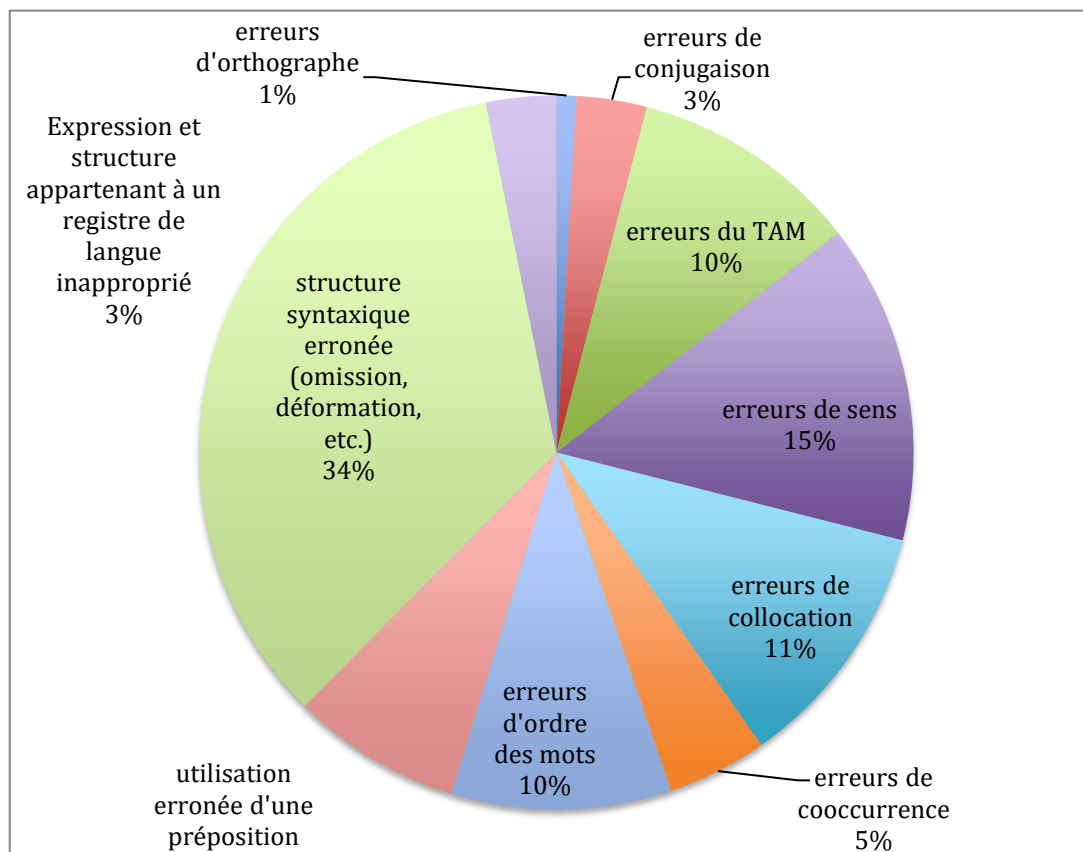


**Figure 5-1 : Répartition des erreurs selon les classes d'erreurs de la typologie**

De manière générale, la plupart des erreurs relevées sont relatives aux propriétés sémantique et syntaxique des verbes (83%), avec en particulier 52% pour celles relevant du niveau de la grammaire et de la syntaxe.

Les erreurs relatives au registre sont les moins fréquentes (3%). Ce nombre peut s'expliquer par le fait que les étudiants chinois apprennent beaucoup de vocabulaire du français standard, notamment dans les manuels de FLE qui utilisent également le registre soutenu. De ce fait, ils connaissent moins bien le langage du registre familier, n'étant pas capables d'établir la distinction entre le registre familier et le registre standard de même, il leur est difficile parfois de faire la distinction entre un usage oral et un usage écrit. Par ailleurs, conformément à notre première observation lors de la constitution de la typologie, les étudiants chinois ayant un niveau B1 acquis font moins d'erreurs au niveau morphologique (14%).

La répartition des erreurs par types est illustrée dans la figure 5-2 :



**Figure 5-2 : Répartition des erreurs selon les types d'erreurs de la typologie**

On peut constater que dans la figure 5-2, la structure syntaxique erronée constitue le type d'erreur le plus fréquent et représente 34% des cas, alors que les erreurs de sens arrivent en deuxième position (15%), suivies par les erreurs de collocation (11%). Par ailleurs, si l'on regroupe les erreurs de collocation et celles de co-occurrence sous la même étiquette (les erreurs de combinatoire lexicale), on obtient un total de 15%. Cette présence non négligeable semble confirmer le constat de nombreuses études mentionnées dans notre état de l'art (Nesselhauf, 2005 ; Bolly, 2011), selon lesquelles ce type d'erreurs constitue des déficits « résistants » propres aux apprenants non-natifs. D'ailleurs, on pourrait penser que les apprenants chinois ne font pas beaucoup d'erreurs sur les collocations et les formes figées en général, mais on verra qu'en fait ils ne les utilisent que très peu par rapport aux natifs (voir la partie 5.3.).

Au niveau morphologique, les erreurs d'orthographe (1%) et de conjugaison (3%) sont les moins fréquentes, alors que la grande majorité des erreurs est liée au TAM (10%).

Observons maintenant les différents types d'erreurs à l'aide d'exemples d'erreurs.

### 5.2.3.1. Au niveau morphologique

#### Erreurs d'orthographe

Nous avons relevé deux erreurs d'orthographe relatives à l'omission de lettres :

1. Généralement dit, l'intertextualité **monte** la relation entre deux ou plusieurs textes. (traduction\_8) (Correction proposée : De manière générale, l'intertextualité montre la relation...)
2. C'est selon les cas concrets qu'il **choit** le mode le plus adéquat. (traduction\_4) (Correction proposée : choisit)

Ces cas très rares pourraient s'expliquer par la négligence des étudiants et on pourrait même dire que ce sont des coquilles, des erreurs de frappe et pas de méconnaissance. Ainsi, ces erreurs d'orthographe ne semblent pas représenter des difficultés.

#### Erreurs de conjugaison

Parmi les sept erreurs de conjugaison relevées, trois cas se rapportent au verbe écrit à l'infinitif là où un mode et un temps serait à utiliser, comme le montre l'exemple 3 :

3. Influencé par ces autres, plutôt par leurs pensées et paroles, l'auteur trouve le nouveau sens dans ces anciennes pensées et paroles et **proposer** ses propres idées. (traduction\_8) (Correction proposée : et propose ses propres idées)

Comme on peut le constater, la phrase comporte deux verbes et nous pouvons supposer que l'étudiant a oublié de conjuguer le deuxième verbe *proposer*. Par ailleurs, nous avons également identifié trois cas ayant trait aux temps composés du passé. Il s'agit d'erreurs sur l'accord de l'auxiliaire ou l'accord du participe passé :

4. Les quatre principes « L'égalité et des avantages mutuels, mettant l'accent sur des résultats concrets, la diversité et le développement commun » **est proposé** dans cette période.



(socioculturel\_1) (Correction proposée : sont proposés)

On relève souvent cette erreur chez les natifs aussi car ils oublient quel est leur sujet grammatical. Ainsi, l'origine de ces erreurs pourrait être due à la négligence. Notons par ailleurs que les verbes du LST étudiés sont majoritairement des verbes réguliers et leur conjugaison n'est pas compliquée pour les étudiants chinois.

*Erreurs de temps/d'aspect/de modalité des verbes (TAM)*

Comme souligné plus haut, la plupart des erreurs morphologiques sont liées au TAM. Dans la plupart du temps, les apprenants ont employé le temps du passé simple de l'indicatif plutôt que le présent ou le participe passé, comme dans l'exemple suivant :

5. Dans ses plusieurs ouvrages, il nous **montra** que le capitalisme se développe grâce à l'existence d'une catégorie sociale. (Correction proposée ; Dans plusieurs de ses ouvrages, il nous a montré) (littérature\_6)

Dans la phrase *il nous montra* qui implique un lien avec le présent, au lieu d'utiliser le passé composé, l'étudiant a choisi le passé simple. Il faut noter que le passé simple, souvent utilisé dans le récit écrit, est quasiment absent dans l'écrit scientifique. Les étudiants chinois ont souvent rencontré le temps du passé simple dans des œuvres littéraires sur lesquelles ils travaillent souvent et qu'ils ont très peu travaillé avec des textes scientifiques. Par conséquent, ils sont peu familiarisés avec les temps présents dans le genre de l'écrit scientifique, comme en témoigne l'exemple 6 où l'étudiant a utilisé à la fois le futur proche et le futur simple :

6. Ainsi, nous **irons expliquer** les différences culturelles sino-françaises dans la traduction dans les aspects de la culture écologique... (traduction\_9) (Correction proposée : nous expliquerons...)

Il semble, dans notre corpus, que les étudiants chinois connaissent les temps du verbe mais qu'ils ne maîtrisent pas les nuances de leurs valeurs sémantiques qui donnent lieu à des erreurs comme celles présentées ci-dessus.

### 5.2.3.2. Au niveau sémantique et des co-occurrences lexicales

#### Erreurs de sens

Comme nous l'avons souligné plus haut, les erreurs de sens occupent le deuxième rang (15%). Parmi les 32 erreurs recensées, la plupart des erreurs viennent d'une confusion entre deux verbes synonymes :

7. Sa lecture de Freud ne se limite pas à l'orthodoxie freudienne, il veut saisir les pensées d'origine de Freud, c'est à dire de **montrer** Freud à la freudienne. (littérature\_8) (Correction proposée : c'est à dire présenter Freud 'à la freudienne')
8. Il reçoit les signaux intertextuels dans le texte original et essayer de le **montrer** dans le texte traduit d'une façon compréhensible pour le lecteur de la version traduite. (traduction\_8) (Correction proposée : Il décrypte les signaux intertextuels dans le texte original et essaie de les restituer dans le texte traduit...)
9. Cette œuvre se compose totalement de trois parties, la première partie **note** en désordre le souvenir et la méditation... (Correction proposée : la première partie décrit pêle-mêle le souvenir...) (littérature\_4)

Il est intéressant de remarquer que ce sont notamment les verbes *noter* et *montrer* qui posent problème, car ils sont courants et donc mieux connus des étudiants chinois qui en font un suremploi. Dans l'exemple 7, le verbe *montrer* est utilisé pour signifier « porter quelque chose à la connaissance de quelqu'un » plutôt que « démontrer, mettre en évidence » parce que la phrase n'implique pas un processus de justification. Quant à l'exemple 8, le lien sémantique entre le verbe *montrer* et *restituer* est moins évident, mais ils consistent à rendre quelque chose visible. L'étudiant ne semble pas totalement maîtriser le sens du verbe *montrer* (*faire voir, mettre en évidence quelque chose tel qu'il est*), puisque dans ce contexte de traduction, il s'agit de « rétablir » les signaux intertextuels dans son état original. Dans l'exemple 9, il existe un lien sémantique entre les verbes *noter* et *décrire* qui se rapportent à l'écrit. L'emploi erroné du verbe *noter* nous permet de voir que l'étudiant n'a pas bien saisi le sens de ce verbe : *noter* signifie « mettre une observation par écrit » ou « remarquer » au sens figuré, alors qu'il s'agit de « représenter » le souvenir dans l'œuvre.

Pourtant, on doit admettre qu'il n'est pas toujours aisé d'expliquer l'origine de ces erreurs, mais on pourrait supposer que les étudiants chinois qui ont l'habitude de recourir à des dictionnaires sino-français (ceux-ci présentent souvent des traductions littérales sans contexte) ne maîtrisent pas totalement les nuances sémantiques des mots.

### Erreurs de collocation

Les erreurs de collocation représentent 11% (25 cas) des erreurs relevées. Comme nous nous intéressons aux verbes du LST, la majorité des erreurs sont des collocations verbales ayant un nom comme base et un verbe comme collocatif. On peut distinguer deux cas différents : le premier cas concerne l'utilisation d'un nom imprécis alors que le deuxième se rapporte au choix d'un collocatif inapproprié.

Les trois exemples suivants illustrent le premier cas :

10. Cependant, au fur et à mesure du développement de la relation entre la linguistique et la théorie de la traduction, on commence à croire qu'il n'y a que la linguistique qui pourra décrire et **expliquer un flot de choses** dans la traduction. (traduction\_6) (Correction proposée : décrire et expliquer tous les éléments qui interviennent dans la traduction)
11. Au XXe siècle, les dirigeants chinois ont visité fréquemment plusieurs pays africains, ainsi que la Chine a organisé plusieurs fois les sommets et les forums sino- africains, **tout montre** l'attention chinoise au continent africain. (socioculturel\_1) (Correction proposée : et de même la Chine a-t-elle organisé plusieurs fois les sommets..., tout ceci montre une attention de la Chine pour l'Afrique)

Comme on peut le constater, les noms objets imprécis sont des noms « passe-partout » comme *tout*, *chose*. Dans l'exemple 10, on peut supposer que l'étudiant a une connaissance limitée sur les noms d'objets qui peuvent apparaître en co-occurrence avec le verbe *expliquer*, ce qui l'amène à employer un terme vague et imprécis. Quant à l'exemple 11, il semble que l'étudiant a voulu utiliser la locution *tout ceci montre*, mais il ne maîtrise pas bien son insertion syntagmatique. Le deuxième cas d'erreur de collocation rencontré est l'emploi d'un collocatif inapproprié (voir les exemples suivants) :

12. Comme un grand maître en roman social et psychologique, François Mauriac **choisit** aussi **des problèmes** réalistes qui existaient dans les familles les plus communes de la société capitaliste...(littérature\_6) (Correction proposée : s'est intéressé aux problèmes...les plus représentatives)
13. En fait, le développement de l'histoire de l'éducation moderne est l'évolution de l'idée éducative moderne. On peut conclure (sur) les trois caractères ci-dessous en **étudiant ces pensées** pour mieux faire la réforme de nos jours. (collocations: en nous inspirant de ces idées) (socioculturel\_2)
14. Nous essayons quand même de **proposer un peu d'opinions** sur les problèmes discutables dans le domaine de traduction. (traduction\_8) (collocation: Nous essayons également de donner nos opinions sur certains sujets à discussion dans le domaine de traduction.)

Dans l'exemple 12, le choix du collocatif *choisir* ne semble pas correspondre à une collocation typique de la base *problème*. En effet, l'étudiant souhaite préciser que l'écrivain François Mauriac s'intéresse à des problèmes réalistes. Concernant l'exemple 13, le choix du collocatif n'est pas si mauvais, mais si l'on considère le contexte en amont, le nom *pensée* pose problème dans le contexte et le sens de la collocation ne correspond pas au sens attendu. Enfin, dans l'exemple 14, la collocation *proposer des opinions* est quasiment absente dans l'écrit scientifique, d'autant plus qu'en employant *proposer un peu d'opinions*, l'étudiant n'a pas respecté la contrainte imposée par le collocatif qui est incompatible avec le déterminant *un peu de*. En effet, l'opinion n'est pas une quantité, donc la valeur quantitative *un peu de* ne peut pas s'appliquer ici.

#### Erreurs de co-occurrence avec les adverbes

En dehors des erreurs de collocations verbales, il existe les erreurs de combinatoire lexicale, notamment [verbe+adverbe]. Au total, dix erreurs ont été relevées. Ainsi, dans l'exemple 15, nous constatons l'emploi d'un néologisme adverbial. Il semble que l'étudiant ait cherché un adverbe qui signifie « d'une façon détaillée » en ajoutant le suffixe *-ment* à l'adjectif *détaillé*. Donc, il connaît la règle de construction des adverbes en /-ement/ mais ne sait pas que celui-ci n'existe pas.

15. En **comparant détaillément** les deux versions, de différents problèmes seront présentés et

analysés. (traduction\_2) (Correction proposée : comparant d'une façon détaillée)

Nous avons également rencontré des cas où les adverbes sont mal choisis et ne correspondent pas au sens exprimé, comme dans les exemples suivants :

16. Il convient de **noter définitivement**, que cette syntaxe courte constitue le trait original du roman...(traduction\_7) (Correction proposée : Enfin, il convient de noter que...)
17. Nous **proposerons grossièrement** une distinction de quatre grands courants textuel, anaphorique, inactuel et aspectuel. (linguistique\_1) (Correction proposée : Nous proposerons sommairement une distinction...)

Dans l'exemple 16, le contexte élargi nous aide à affirmer que l'étudiant a utilisé *définitivement* pour marquer le dernier élément, alors que les adverbes *enfin*,  *finalement* auraient été plus appropriés. Quant à l'exemple 17, le choix de l'adverbe *grossièrement* nous paraît maladroit. Cet adverbe peut signifier 'en employant un langage familier/grossier' ce qui n'entre pas dans le registre attendu dans les écrits scientifiques. Mais il peut signifier aussi 'de façon sommaire, sans détail' et dans ce cas il pourrait être accepté. Toutefois, son caractère polysémique et la prégnance du sens lié à la familiarité des propos, limite son utilisation dans un discours scientifique.

Enfin, les cas relevés ne sont pas nombreux, ce qui peut s'expliquer par le fait que l'emploi des cooccurrences [verbe+adverbe] est restreint dans l'écrit des apprenants. On pourrait penser que les apprenants ne disposent pas des connaissances qui leur permettraient de recourir aux co-occurrences [verbe+adverbe].

Les erreurs présentées ci-dessus nous conduisent à supposer que les étudiants chinois manquent de moyens pour trouver les collocatifs ou les co-occurents appropriés. Les dictionnaires sino-français ne leur proposent pas ces informations de combinatoire lexicale ; ils choisissent par conséquent soit de créer une forme inexistante, soit d'employer un terme qui ne traduit pas le sens visé.

De manière générale, les étudiants chinois n'ont pas une bonne appréhension du phénomène de combinatoire lexicale tant en compréhension qu'en production. Lorsqu'il s'agit de mots inconnus, ils recourent souvent aux dictionnaires sino-français pour vérifier les traductions en français, ce qui pourrait expliquer la

plupart de leurs erreurs à partir de traductions littérales des dictionnaires et des transferts malheureusement souvent erronés du chinois. D'une part, il nous semble important de les sensibiliser au phénomène de collocations/co-occurrences ; d'autre part, on peut suggérer de meilleurs outils ou moyens de consultation qui facilitent l'accès aux informations de combinatoire lexicale.

### 5.2.3.3. Au niveau de la grammaire et de la syntaxe

#### Erreurs d'ordre des mots

Les erreurs d'ordre des mots représentent 10% (22 cas) des erreurs relevées. La plupart des erreurs (17 cas) viennent de la mauvaise place de l'adverbe, comme en témoignent les exemples 18 et 19 :

18. On **souvent considère** *quand* comme connecteur de simultanéité qui peut donner lieu à une lecture consécutive avec l'aspect aoristique. (linguistique\_1) (Correction proposée : on considère souvent)
19. Lawrence Venuti, le linguiste qui est de Italien American (un Italo-Américain), **a proposé d'abord** ces deux stratégies de la traduction cibliste et de la traduction sourcière dans son livre *The Translator's Invisibility*. (traduction\_9) (Correction proposée : a d'abord proposé)

L'origine de l'erreur 18 pourrait être liée à l'influence du chinois où les adverbes se placent toujours devant les verbes. Dans l'exemple 19, l'adverbe *d'abord* n'est pas mis à la bonne place, car c'est souvent à cette place qu'il se trouve à l'oral. Nous avons également relevé quelques cas concernant la place de l'objet indirect dans les temps composés. Par exemple, dans l'exemple 20, l'étudiant a inversé le pronom *nous* et l'auxiliaire *avoir*.

20. Tout au long du roman, l'auteur **a nous décrit** neuf voyages mais pas un seul en apparence. (littérature\_2) (Correction proposée : l'auteur nous a décrit)

#### Utilisation d'une préposition erronée

L'utilisation d'une mauvaise préposition représente 8% (17 cas) des erreurs. La plupart des emplois erronés (7 cas) des prépositions se rencontrent dans les constructions verbales, comme l'illustre l'exemple 21.

21. De plus, c'est peut-être que les étudiants peuvent obtenir leurs diplômes en **choisissant passer** l'examen ou la thèse ou tous les deux... (socioculturel\_2) (Correction proposée : il se peut que les étudiants peuvent...en choisissant de passer...)

Cette erreur vient du non-respect de certaines constructions syntaxiques. On voit que les apprenants n'ont pas bien maîtrisé la construction *choisir de faire quelque chose*. Ici, la préposition contribue à construire le sens de la construction verbale. Le choix de cette préposition est peu prédictible et nous pensons que l'apprentissage de ce type de préposition n'est possible qu'à travers l'acquisition des constructions verbales.

Une autre partie des erreurs (10 cas) se rapporte au sens de la préposition. Par exemple, dans l'exemple 22, l'apprenant aurait dû employer la préposition *dans* qui introduit une localisation alors que dans l'exemple 23, il y a lieu d'utiliser une préposition pour exprimer le but.

22. Mais quelle que ce soit, la version doit être acceptée par le lecteur, au contenu comme à la forme. **Nous l'expliquons par** la paragraphe suivante... (traduction 5) (Correction proposée : tant au niveau du contenu que de la forme... Nous l'expliquons au paragraphe suivant)
23. Nous ne pouvons pas imposer le concept culturel de langue originale au lecteur cible, ne pouvons pas plus utiliser le type et la standard de culture original **à expliquer** les autres cultures (traduction\_9) (Correction proposée : ne pouvons non plus utiliser...pour expliquer)

### Structure syntaxique erronée

Comme nous l'avons souligné, ce type d'erreurs totalise 34% (76 cas) des erreurs relevées, et constitue donc le cas le plus fréquent de notre corpus. Nous pouvons distinguer trois catégories : l'omission d'un élément dans la phrase (12 cas), des structures déformées ou inexistantes (58 cas), ou encore des structures maladroitement (6 cas).

Les cas d'omission peuvent être attribuables soit à la négligence des apprenants soit à une maîtrise insuffisante des constructions syntaxiques. Dans l'exemple 24, il manque le pronom *le* devant le verbe *expliquer*, qui renvoie à la citation entre guillemets. Dans l'exemple 25, l'apprenant aurait dû ajouter l'auxiliaire *être* pour

construire la phrase passive.

24. **Comme explique l'auteur** : « l'élément psychologique, comme l'élément pictural, se libère insensiblement de l'objet avec lequel il faisait corps... » (littérature\_5) (Correction proposée : Comme l'explique l'auteur)
25. Monsieur LUO Dagang qui insiste sur ce que les œuvres littéraires peuvent **considérées comme** le miroir de la vie des écrivains. (Correction proposée : Luo Dagang insiste sur...peuvent être considérées comme...) (littérature\_8)

Au-delà des erreurs sus-citées, la plupart des erreurs (58 cas) de structure syntaxique erronée correspondent à des cas où la structure syntaxique n'est pas respectée. Dans l'exemple 26, la structure est déformée, puisque le sujet est absent et que la construction *pousser à étudier et analyser* est rarement utilisée dans l'écrit scientifique pour établir une relation de cause à effet. Rappelons que dans le chapitre III, nous avons expliqué que les constructions causatives *amener/conduire à faire quelque chose* sont fréquentes dans l'écrit scientifique. Il apparaît qu'ici, l'apprenant manque de moyens linguistiques pour justifier le choix de son sujet portant sur l'étude du passé simple et de l'imparfait.

26. **Après avoir lit la lecture** de quelques livres linguistiques de Gosselin, Touratier, Bres, etc., **nous pousse à étudier et analyser** le passé simple et l'imparfait. (linguistique\_1) (Correction proposée : La lecture des ouvrages de...nous amène à nous intéresser à l'étude et à l'analyse du passé simple et de l'imparfait)

Dans l'exemple 27, il nous semble difficile de trouver le comparant et le comparé puisque la structure syntaxique est déformée. En consultant un contexte plus large, nous avons compris que l'auteur compare deux versions du mythe de Narcisse. Ici, le verbe *comparer* aurait été employé au passif et le sujet *les deux dernières* aurait été mis au début de la proposition. Nous croyons que cette erreur découle de la non maîtrise de la structure actancielle régie par le verbe *comparer*.

27. **Comparant à** la variante la plus populaire ci-dessus citée, même si les deux dernières sont très différentes, il existe encore des points communs. (littérature\_3) (Correction proposée : Comparées à la variante la plus populaire citée ci-dessus, les deux dernières sont très différentes, mais il existe...)



Ces erreurs sont problématiques, car on a du mal à comprendre ce que l'apprenant souhaite exprimer. Par ailleurs, nous pouvons souligner que la maîtrise insuffisante des structures syntaxiques dans l'écrit scientifique conduit souvent l'apprenant à construire des phrases maladroitement (6 cas), comme le montrent les exemples 28, 29 et 30 :

28. **Il est hors de doute que Meschonnic considère** la poétique comme une théorie plus profonde et plus complète que la linguistique. (Correction proposée : Meschonnic considère...) (traduction\_1)
29. Pour traduire ces références et ces noms propres, les traducteurs de nos jours emploient généralement deux façons de traduire : la traduction phonétique (la traduction littérale) et la traduction sémantique. **Voyons le tableau que nous avons fait pour illustrer quelques façons choisies** par les traducteurs. (traduction\_1) (Correction proposée : Nous illustrons dans le tableau suivant les différentes façons de traductions adoptées par les traducteurs)
30. **Il nous vaut mieux de considérer** le paysage de la littérature comme une combinaison de la littérature et le paysage, au lieu de considérer la littérature comme un miroir solitaire qui ne reflète ou déforme que le monde extérieur. (littérature\_7) (Correction proposée : il nous semble préférable de considérer...)

On voit bien que ces exemples sont beaucoup plus compréhensibles que l'exemple 27. Les exemples 28 et 29 peuvent être considérés comme une maladresse, car ils ne sont pas totalement faux du point de vue strictement linguistique. Ici, il nous semble intéressant de les analyser avec l'exemple 30 qui est incorrect parce que ces trois exemples font bien ressortir le fait que les apprenants ne disposent pas de moyens linguistiques pour exprimer certaines fonctions rhétoriques dans leur écrit académique. Il s'agit de proposer une illustration avec le tableau (l'exemple 29), de référer au propos d'un chercheur (l'exemple 28) et d'introduire un point de vue (l'exemple 30). Nous croyons que ces erreurs ou plutôt, maladroites, sont attribuables au fait que les apprenants sont peu familiarisés avec le genre de l'écrit scientifique ainsi que les constructions syntaxiques spécifiques ou fréquentes dans l'écrit scientifique, trahissant par conséquent le statut de novices des apprenants chinois dans le domaine de l'écrit scientifique.

#### 5.2.3.4. Au niveau du registre

##### Expression et structure appartenant à un registre inapproprié

Les sept erreurs relevées (3%) se rapportent au cas où l'apprenant emploie une expression ou une structure appartenant à un registre inapproprié, en mélangeant notamment des usages oraux et écrits, comme dans les deux exemples suivants :

31. Quand Simon parle du sentiment de la création de La Route des Flandres, **il explique comme ça** : toutes les choses surgissent ensemble dans mon esprit... (littérature\_12)  
(Correction proposée : il donne les explications suivantes)
32. **Ce qu'on doit étudier, c'est que** dans un modèle ou une circonstance si centralisée, comment on a un pluralisme d'enseignement, le système des différents niveaux du grade. (socioculturel\_2) (Correction proposée : Dans un modèle..., il nous semble important d'étudier comment on peut avoir un enseignement pluraliste)

Les erreurs relatives au registre sont parmi les moins fréquentes, mais elles n'ont pas le même statut que les erreurs de type morphologique. En effet, nous pensons que ces erreurs relatives au registre sont plus problématiques, car elles ne respectent ni la norme linguistique, ni les codes linguistiques reconnus par la communauté de discours. L'expression *comme ça* et la structure *ce que...c'est que* apparaissent à l'oral mais ne sont pas acceptables à l'écrit. Ce type d'erreurs peut être attribuable au fait que l'apprenant n'est pas conscient de l'existence des usages oraux et écrits de la langue. Nous pensons que le fait de sensibiliser l'apprenant aux constructions verbales fréquentes et à la notion d'usage dans l'écrit scientifique, en les distinguant des autres formes à la fois d'usage et de registres, permet de réduire ce type d'erreur.

#### 5.2.3.5. Réflexions sur l'analyse des erreurs

À l'issue de l'analyse des erreurs liées à l'utilisation des verbes du LST et de leurs constructions verbales, nous avons pu identifier les difficultés des apprenants qui se situent notamment aux niveaux sémantique et syntaxique (83%). Ceci concerne en particulier la structure syntaxique erronée (34%), les erreurs de combinatoire lexicale (15%) et de sens (15%). Il est intéressant de noter que nous avons pu relever certaines erreurs révélant les lacunes linguistiques des apprenants chinois pour

exprimer certaines fonctions rhétoriques dans leur écrit académique. Le nombre d'erreurs élevé en matière de structures syntaxiques a mis en lumière l'importance d'enseigner les constructions verbales et d'aider l'apprenant à comprendre les constructions dans le discours scientifique.

Les sources possibles des erreurs relevées sont assez diverses : une surcharge cognitive (négligence), l'influence de la langue maternelle, le manque de savoir sur les propriétés sémantique et syntaxique des verbes français, le manque de connaissances sur les codes linguistiques dans l'écrit scientifique. Parmi ces causes, dans le cadre de cette thèse, les deux dernières paraissent primordiales. En effet, l'analyse du comportement lexico-syntaxique des verbes du LST dans le chapitre IV nous permet d'envisager des remédiations. Il nous semble que le fait de sensibiliser les apprenants à la phraséologie associée au verbe (par exemple, collocations, constructions syntaxiques, routines verbales) pourra les aider à mieux saisir les sens du verbe et à comprendre le fonctionnement de la langue scientifique. Ce processus de sensibilisation pourrait être développé dans des activités d'observation sur le corpus que nous expliquerons plus en détail dans la section 5.5.

De manière générale, d'une part, on constate que les apprenants ne maîtrisent pas bien la langue, car ils sont encore en phase d'acquisition de la compétence langagière ; d'autre part, si l'on considère l'emploi de temps verbaux inappropriés et la production de constructions maladroites, on peut supposer que les apprenants connaissent insuffisamment les codes linguistiques de l'écrit scientifique. En fait, à notre connaissance, le genre de l'écrit scientifique ne fait pas l'objet d'un enseignement spécifique pour ce public. On pourrait faire l'hypothèse que certaines lacunes liées à cette dimension empêchent l'apprenant d'employer les expressions et les constructions appropriées au genre et au contexte. C'est pourquoi nous pensons que ces lacunes pourraient être notamment mises en évidence par l'analyse des cas de sous-emploi et de suremploi en comparaison avec les écrits de natifs.

### **5.3. Diagnostic II : spécificités liées aux constructions verbales du LST**

Notre diagnostic se décompose de deux parties, se focalisant non seulement sur les erreurs liées aux constructions verbales mais aussi sur les cas de suremploi et de sous-emploi de verbes du LST et de constructions verbales. Cette dernière approche, peu utilisée dans le domaine du FLE, constitue un point primordial de notre analyse.

Dans cette section, nous tentons de proposer une grille de diagnostic sur les spécificités qui guideront notre analyse comparative entre le Corpus Sup-Chinois et le Corpus Sup-Natifs. Cette grille sert à révéler les verbes du LST et leurs constructions verbales sur-employées ou sous-employées, permettant ainsi d'évaluer dans quelle mesure les apprenants chinois maîtrisent les constructions spécifiques de l'écrit scientifique.

Nous proposerons d'abord un aperçu général des travaux mettant en évidence des cas de suremploi et de sous-emploi dans les corpus d'apprenants. Ensuite, nous présenterons notre grille fondée sur ces recherches avant d'effectuer les analyses comparatives.

#### **5.3.1. Études représentatives sur les spécificités dans le domaine de l'EAP**

Force est de constater qu'en dehors des erreurs orthographiques, lexicales et grammaticales, la langue de l'apprenant se distingue de celle des natifs par des cas de suremploi et de sous-emploi de mots, d'expressions et de structures syntaxiques (*cf.* Chapitre I). Néanmoins, à notre connaissance, il existe peu d'études portant sur ces spécificités en FLE, et encore moins dans le domaine des écrits universitaires. Faute de réels corpus d'apprenants de FLE en France, on ne dispose pas de statistiques de ce genre. Dans la section suivante, nous allons notamment examiner des études en anglais qui peuvent nous donner une idée de ce phénomène en français. Néanmoins, nous pouvons d'ores et déjà mentionner quelques rares études en FLE : l'étude de Bolly (2007 ; 2008) portant sur des séquences verbo-nominales (*ex : prendre*

*connaissance de, donner une brève histoire de*) dans les textes académiques et l'étude de Forsberg (2010) consacrée aux séquences phraséologiques à l'oral (ex : *disons que, c'est vrai que*).

Dans le domaine de l'EAP, ces spécificités sont étudiées de manière plus systématique, notamment sous trois aspects avec la méthode d'analyse Contrastive Interlanguage Analysis (CIA) (*cf.* Chapitre II) (Flowerdew, 2015), l'analyse phraséologique (Hyland, 2008a ; Chen & Baker, 2010), la comparaison par rapport aux listes de l'EAP comme Academic Word List (AWL) (Paquot, 2010) et le métadiscours<sup>3</sup> (Hyland & Milton, 1997 ; Luzón, 2009). Nous allons présenter ci-dessous certaines études représentatives qui abordent la question des spécificités dans les écrits des apprenants sous ces trois aspects.

### 5.3.1.1. Études portant sur l'analyse phraséologique

Dans le chapitre I, nous avons présenté les éléments phraséologiques auxquels appartiennent les patrons verbaux comme des déficits «résistants» chez les apprenants du fait de leur complexité linguistique. Ainsi, la phraséologie est un aspect essentiel permettant de distinguer clairement les productions des natifs de celles des apprenants. (Ebeling & Hasselgård, 2015).

Étant donné que notre intérêt porte sur les constructions verbales, nous présentons deux études consacrées aux constructions verbales dans les écrits des apprenants afin de comprendre ce que font les apprenants avec ces éléments. Nous pourrions ensuite faire des choix quant aux éléments à aborder pour la classe de FLE.

L'étude de Wulff et Gries (2011) porte sur les constructions avec l'infinitif ou le gérondif comme compléments de verbes. Au niveau sémantique, ces deux constructions se distinguent par le fait que la première désigne une possibilité (ex : *I try to finish it. /J'essaie de le finir.*) alors que la deuxième renvoie à un événement réel (*I stop talking. /J'arrête de parler.*). Wulff et Gries ont examiné ces deux

---

<sup>3</sup> Il vient du terme anglais « metadiscourse » traduit par Grossmann (2012). Ce terme est employé par Hyland pour désigner « l'ensemble des marques impliquant une forme de réflexivité du scripteur dans le cadre de la négociation de l'interaction avec le lecteur appartenant à une communauté spécifique ». (*ibid.*, p. 149)

constructions dans un sous-corpus de l'*International Corpus of Learner English* (ICLE), composé d'écrits d'apprenants allemands de niveau avancé, en le comparant à la partie britannique du corpus d'experts *International Corpus of English* (ICE-GB). Leur objectif est d'observer si les apprenants sont conscients de l'association entre les verbes et ces constructions et si les verbes sont utilisés dans les constructions à bon escient. Ils ont extrait dans un premier temps tous les verbes associés à ces deux constructions dans les deux corpus. L'analyse comparative a montré que certains verbes (ex : *to prefer, to continue, to fail, to hope*) fréquemment utilisés dans la construction avec l'infinitif comme complément chez les experts sont davantage employés dans la construction avec le gérondif chez les apprenants, tandis que d'autres verbes (ex : *to begin*) employés de préférence dans la construction avec le gérondif chez les experts apparaissent dans la construction avec l'infinitif chez les apprenants. En plus de ces différences, les auteurs ont souligné que certains verbes fréquemment associés à ces deux constructions sont absents chez les apprenants. Ces verbes doivent donc faire l'objet d'un enseignement systématique, par exemple, *to seek* et *to continue* utilisés dans la construction avec l'infinitif et *to envisage, to fancy* et *to imagine* dans la construction avec le gérondif (*ibid.*, p. 80). Ainsi, les auteurs ont conclu sur l'importance de l'E/A des associations entre les verbes et les constructions en insistant sur le fait que les apprenants doivent comprendre le sens véhiculé par ces associations pour améliorer l'exactitude de leurs productions écrites.

Toujours dans une perspective phraséologique, Gilquin (2012) s'attache à étudier les constructions causatives en anglais. Par exemple, la construction causative comportant le verbe *to make* :

- (1) X MAKE Y Vinf : *This does not make the scam legitimate.* (traduction : *Ceci ne rend pas l'escroquerie plus légitime.*)
- (2) X MAKE Y Vpp : *The effects are making themselves felt.* (traduction : *Les conséquences se font sentir.*)

En se basant sur le corpus BNC, Gilquin a montré que ces constructions causatives s'associent de préférence à certains verbes. Par exemple, pour la construction causative X MAKE Y Vpp, les verbes *known, felt* et *understood* représentent 83% des occurrences. Les auteurs comparent ensuite ces résultats aux

données du corpus d'apprenants ICLE et relèvent ainsi des différences d'utilisation à plusieurs niveaux : 1) erreurs syntaxiques comme *make us not to forget* (<ICLE-JP>), 2) le suremploi et le sous-emploi de certaines constructions causatives (ex : la construction X MAKE Y Vinf est sur-employée par des apprenants chinois), 3) par rapport aux experts, l'emploi des constructions causatives chez les apprenants est « unidiomatic » (non idiomatique) parce qu'ils utilisent des verbes variés et n'ont pas tendance à utiliser les constructions préférées des experts. Par exemple, le verbe *become* présent chez les apprenants dans la construction causative *make women become aware of their rights* (<ICLE-RU>), est absent chez les experts.

Par ailleurs, les auteurs ont suggéré que ces différences peuvent être expliquées par trois éléments tels que le manque de connaissances, le transfert de la langue maternelle et le manque de matériels d'enseignement appropriés.

Enfin, de manière générale, on peut dégager de ces deux études trois caractéristiques principales liées à l'utilisation des constructions verbales chez les apprenants :

1. l'association faible entre le verbe et ses constructions
2. l'emploi de constructions présentant une variété moins riche que celles des natifs
3. l'emploi de constructions à tendance non idiomatique.

Ces trois caractéristiques seront prises en compte lors de notre diagnostic sur les spécificités.

### 5.3.1.2. Études portant sur les listes de l'EAP

La constitution des listes de l'EAP telles que AWL et AKL (*cf.* Chapitre I) a conduit les chercheurs à examiner si les mots considérés comme essentiels dans l'EAP sont présents dans les écrits des apprenants, et si oui, comment ces derniers les utilisent par rapport aux experts. La comparaison des corpus basée sur ces listes constitue ainsi un autre aspect permettant de mettre en évidence les spécificités des apprenants.

Basée sur la liste de l'AWL, Chen et *al.* (2009) ont calculé la fréquence des mots

appartenant à cette liste dans les résumés des revues relevant des domaines du commerce et du management. Les auteurs ont mis en contraste des revues publiées à Taiwan et celles publiées à l'international. De manière générale, les résultats ont montré que les mots de l'EAP sont plus présents dans les revues publiées à l'international. Ainsi, les auteurs soulignent l'importance de ces mots pour les étudiants s'ils souhaitent améliorer leurs compétences en lecture et en écriture, en se rapprochant des normes internationales.

De même que Chen et *al.*, Paquot (2010) s'intéresse aux mots essentiels dans l'EAP. Son étude se fonde sur l'élaboration de la liste de l'AKL et comprend une dimension rhétorique dans la mesure où elle se focalise sur les mots remplissant des fonctions rhétoriques importantes telles que l'expression de l'effet et de la cause, l'illustration par des exemples, la comparaison, etc. En comparant le corpus ICLE à un sous-corpus de BNC, l'auteure a souligné qu'environ 50% de mots de l'AKL sont sous-employés par les apprenants dont les noms comme *argument*, *difference* et les verbes tels que *argue*, *illustrate*, *exemplify* et *explain*. Il a été également montré que les apprenants ont un répertoire lexical plus limité au sujet du choix de mots ou d'expressions revêtant certaines fonctions rhétoriques. Par exemple, ils se limitent aux adverbiaux *for example* et *for instance* pour exemplifier, ignorant les mots spécifiques remplissant cette fonction tels que *illustrate* et *exemplify*. Les fonctions rhétoriques constituent donc un aspect non négligeable dans l'étude sur les spécificités que nous allons approfondir dans la section suivante.

### 5.3.1.3. Études portant sur le métadiscours

De nombreux travaux (Hyland & Milton, 1997 ; Ädel, 2003 ; Hyland, 2004 ; Noble, 2006) tendent à montrer que les difficultés rencontrées par les apprenants de niveau avancé, se situent essentiellement au plan des fonctions communicatives et pragmatiques : des difficultés à interagir avec le lecteur, à se référer au discours d'autrui, à atténuer l'assertion, à exprimer son point de vue ou à structurer le discours. Précisons que la question du métadiscours concerne différents types de marqueurs dont font partie les constructions verbales.

Arrêtons-nous sur deux études du domaine : 1) celle de Neff et *al.* (2004) portant



sur le positionnement de l'auteur et 2) celle de Hyland (2008a) qui examine différentes fonctions rhétoriques (interaction, organisation du discours, l'assertion) avec une approche phraséologique.

Neff et *al.* (2004) ont examiné les marqueurs de positionnement (ex : *I think, it's true that*) dans trois corpus différents : corpus d'apprenants d'anglais composé d'écrits argumentatifs produits par un public hétérogène (des apprenants espagnols, italiens, français et allemands), corpus d'étudiants états-uniens (écrits argumentatifs) et corpus d'articles de recherche d'experts. À l'issue de l'étude contrastive, les auteurs ont constaté dans le corpus d'apprenants, d'une part, un suremploi du pronom personnel *I* associé aux verbes d'opinion comme *think, believe* et des constructions servant à exprimer l'évidence du fait (ex : *it's clear that, it's well known that*) ; et d'autre part, un sous-emploi des phrases modalisées et un manque de justification de l'assertion (en particulier après l'assertion certaine de l'auteur). Par ailleurs, il s'avère que les étudiants états-uniens occupent une place intermédiaire entre les apprenants non-natifs et les experts. Par exemple, comme les apprenants non-natifs, ils utilisent davantage de constructions servant à exprimer l'évidence, ce qui révèle qu'ils ont des difficultés à se positionner et à remettre en cause ce qu'ils lisent, trahissant ainsi leur statut de novice.

Avec une approche phraséologique, Hyland (2008a) a examiné des segments répétés à 4 mots dans un corpus de mémoires et de thèses produits par des apprenants d'anglais de HongKong venant de quatre disciplines : électrotechnique, commerce, linguistique appliquée et microbiologie. Il l'a comparé à un corpus d'articles de recherche d'experts des mêmes domaines. De manière générale, il a montré que par rapport aux experts, les apprenants (en particulier ceux en master) sous-utilisent les expressions en matière d'organisation du discours (*on the other hand, as a result of*) et celles ayant une fonction interactive (*as can be seen, as shown in figure, should be noted that*). En outre, cette étude indique que les écrits des apprenants se caractérisent par des éléments phraséologiques moins variés que ceux des experts.

Ayant mis en évidence les lacunes au sujet des fonctions rhétoriques dans le discours scientifique, ces deux études s'accordent à conclure à la nécessité de l'E/A de ces marqueurs lors de la rédaction d'écrits universitaires par des apprenants de

langues étrangères à l'université. Enfin, il nous semble donc pertinent de relier les constructions verbales à leurs fonctions rhétoriques lors de notre diagnostic afin de mettre en avant les lacunes au niveau discursif.

Nous avons présenté certaines études représentatives qui abordent la question des spécificités sous trois aspects différents (phraséologie, listes de l'EAP et métadiscours). Comme le souligne Granger (2009, p. 19, traduit par l'auteur), ces travaux contribuent à « identifier les caractéristiques lexicales, grammaticales et discursives qui distinguent les productions des apprenants de la norme ciblée ».

Néanmoins, précisons que les trois aspects mettant en avant les spécificités des apprenants sont étroitement liés. De notre point de vue, ces études peuvent donner un aperçu plus complet sur les lacunes des apprenants. L'étude de Granger et Paquot (2009b), présentée ci-dessous, prend en compte ces trois aspects.

#### **5.3.1.4. Études de Granger et Paquot portant sur les verbes de l'EAP**

L'étude de Granger et Paquot (*ibid.*) nous intéresse particulièrement parce qu'elle est très proche de la nôtre dans la mesure où son objectif est d'examiner l'emploi des verbes de l'EAP (les verbes en anglais qui remplissent presque le même rôle que les verbes du LST) (*cf.* Chapitre I) dans un corpus d'apprenants, et ceci par rapport aux corpus d'experts et à celui des étudiants natifs.

Dans un premier temps, les auteures ont examiné respectivement les 100 verbes les plus fréquents dans le corpus d'apprenants ICLE et un corpus d'experts. À travers la comparaison, il s'est avéré que 55,4% de ces verbes sont sur-employés tandis que 29% sont sous-employés par les apprenants. Les auteures ont remarqué que parmi les verbes sous-employés, 88% appartiennent à la liste de l'AKL dont les verbes *describe*, *suggest*, *note*, *include*. Quant aux verbes sur-employés, 90% relèvent de l'anglais langue générale dont les verbes *want*, *feel*, *dream*. Les premiers résultats ont donc conduit à dégager des lacunes liées aux verbes de l'EAP.

Dans un deuxième temps, les auteures se sont focalisées sur les différences sur les formes verbales afin d'examiner l'emploi des patrons verbaux comportant ces formes chez les apprenants. Les formes verbales étudiées comprennent le temps au

passé (*drank*), le participe présent (*drinking*)/passé (*drunk*), l’infinitif (*to drink*) et la forme -s (*drinks*). De manière générale, les résultats ont révélé d’un côté, la sous-exploitation de certaines formes verbales comme le temps au passé et le participe passé (*showed, argued*) et de l’autre côté, le suremploi de la forme à l’infinitif (*to conclude*).

Ensuite, les auteures ont interprété ces résultats en faisant un lien avec les fonctions rhétoriques, affirmant que les formes verbales sous-employées indiquent la non maîtrise des patrons servant à certaines fonctions rhétoriques qui sont pourtant au cœur du genre de l’écrit académique. Il s’agit par exemple du patron ‘as VVN in/by’, utilisé pour se référer à d’autres travaux (*as suggested by*) ou pour renvoyer aux éléments péri-textuels (*as shown in*).

Enfin, en comparant le sous-corpus d’apprenants français (ICLE-FR) à un corpus d’étudiants états-uniens, les auteures ont souligné que certaines difficultés concernant l’emploi des patrons spécifiques à l’écrit académique (ex : le sous-emploi des constructions au passif) se rencontrent également chez les étudiants natifs et que ces derniers occupent ainsi une place intermédiaire entre les apprenants et les experts. En revanche, il a été montré que les apprenants sont confrontés à plus de difficultés en ce qu’ils utilisent beaucoup de collocations et de constructions erronées et « non standards », notamment du fait du transfert de la langue maternelle.

S’inscrivant dans la lignée de cette étude, nous souhaitons proposer un diagnostic sur les spécificités en deux étapes : les verbes du LST, les constructions verbales et les fonctions associées, mettant en évidence les difficultés des apprenants aux niveaux lexical, syntaxique et discursif. Néanmoins, il est important de souligner que contrairement au corpus ICLE, qui rassemble de courts essais d’étudiants, notre corpus d’apprenants contient de longs écrits académiques.

### **5.3.2. Constitution de la grille de diagnostic sur les spécificités**

Cette section sera consacrée à la constitution d’une grille de diagnostic mettant en avant les spécificités liées à l’utilisation des verbes du LST et de leurs constructions verbales dans les écrits d’apprenants chinois. Cette grille va guider

notre analyse comparative entre le Corpus Sup-Chinois et le Corpus Sup-Natifs. Les données de ces deux corpus seront confrontées aux résultats obtenus du corpus d'experts Scientext-SHS.

Nous retiendrons alors trois critères de reconnaissance pour notre comparaison :

- (1) D'abord, les cas de suremploi et de sous-emploi reflètent la non-maîtrise du genre de l'écrit scientifique par les apprenants. Comme le soulignent Granger et Paquot (2009b), ces spécificités constituent un handicap certain pour les apprenants, qui ne peuvent alors exprimer leur pensée avec nuance et dans le style approprié.
- (2) Ensuite, les spécificités liées aux écrits universitaires des apprenants sont peu prises en compte dans l'enseignement du FLE, ce qui peut s'expliquer par le fait qu'elles sont plus difficiles à détecter par rapport aux erreurs et qu'elles demandent un travail de constitution de corpus complexe.
- (3) Enfin, à notre connaissance, les spécificités liées aux constructions verbales dans les écrits scientifiques des apprenants sont encore peu explorées dans le domaine de l'EAP comme dans le FLE. Les chercheurs se sont interrogés notamment sur certaines constructions verbales particulières (ex : constructions causatives) (Wulff & Gries, 2011 ; Gilquin, 2012) mais pas toutes et peu encore pour le français. Notre étude se distingue de ces travaux par le fait qu'avec une entrée verbale, nous souhaitons relever les principales difficultés des apprenants liées à l'utilisation des verbes du LST, notamment en termes de constructions syntaxiques.

La constitution de la grille se fonde sur les analyses linguistiques des patrons verbaux dans le corpus Scientext-SHS qui nous révèlent les principales propriétés sémantiques et syntaxiques des verbes du LST dans l'écrit scientifique (*cf.* Chapitre IV). Rappelons par ailleurs qu'une étude effectuée précédemment (Hatier & Yan, 2017) nous a également permis d'extraire les difficultés des apprenants chinois. Dans une perspective didactique, nous pouvons formuler les hypothèses suivantes :

1. Les apprenants chinois sous-emploient les verbes du LST par rapport aux novices français et aux experts.
2. Les apprenants chinois sous-emploient les constructions verbales qui remplissent souvent des fonctions rhétoriques importantes (ex : exprimer son point de vue, apporter une justification ou un argument).
3. Les apprenants chinois suremployent les mots, les expressions et les constructions non conformes au discours scientifique.

Nous présentons la grille de diagnostic dans le tableau ci-dessous :

Paramètre	Processus
Fréquence des verbes	<ol style="list-style-type: none"> <li>1. Observer les 20 verbes du LST les plus fréquents chez les experts et répartition chez les apprenants chinois et les novices français</li> <li>2. Observer les 20 verbes du LST les plus fréquents chez les apprenants chinois (par rapport aux experts) et répartition chez les novices français</li> <li>3. Observer les 20 verbes du LST les plus fréquents chez les novices français (par rapport aux experts) et répartition chez les apprenants chinois</li> </ol>
Fréquences des constructions verbales	<ol style="list-style-type: none"> <li>1. Observer la fréquence des constructions verbales récurrentes pour les 10 verbes fréquents étudiés (verbes d'analyse et verbes de positionnement) chez les apprenants chinois et les novices français</li> <li>2. Observer la diversité de ces constructions verbales chez les apprenants chinois et les novices français</li> <li>3. Observer, pour les 10 verbes, les constructions suremployées par les apprenants chinois</li> </ol>

**Tableau 5.4 : Grille de diagnostic sur les spécificités**

Comme on peut le constater, la grille présente trois caractéristiques principales :

- Elle est fondée sur les analyses des verbes (verbes d'analyse et verbes de positionnement) ;
- Elle est constituée de deux étapes d'analyse : la comparaison par rapport à la liste du LST et aux constructions verbales ainsi que l'observation des fonctions rhétoriques associées.

Nous pouvons désormais aborder l'étude comparative des corpus.

### 5.3.3. Analyse des spécificités

Lors de la comparaison des corpus, nous nous focalisons en particulier sur le comportement lexico-syntaxique des verbes du LST dans les mémoires des étudiants, à travers les analyses quantitative et qualitative.

Dans un premier temps, nous évaluons le nombre des verbes du LST (342 verbes extraits du corpus Scientext-SHS) dans le corpus d'apprenants chinois et le corpus de novices français afin de voir quels sont les verbes surreprésentés et sous-représentés dans les écrits d'apprenants chinois (par rapport aux novices français et aux experts). Dans un deuxième temps, nous reprenons les 10 verbes retenus pour le diagnostic des erreurs (*choisir, considérer, comparer, décrire, étudier, expliquer, montrer, noter, proposer, supposer*) afin d'examiner leurs constructions verbales présentes dans les deux corpus. Celles-ci seront ensuite comparées aux patrons verbaux obtenus dans le chapitre IV. Finalement, pour les 10 verbes retenus, nous observons, chez les apprenants chinois et les novices français, la fréquence et la diversité des routines verbales remplissant les fonctions rhétoriques.

Notons par ailleurs que dans l'analyse des spécificités sur les 10 verbes, nous n'abordons pas les fréquences brutes du corpus Scientext-SHS puisqu'au niveau quantitatif, ce dernier n'est pas comparable aux deux autres corpus.

#### 5.3.3.1. Fréquences des verbes du LST

Avant de procéder à une analyse qualitative pour les 10 verbes du LST (*choisir, considérer, comparer, décrire, étudier, expliquer, montrer, noter, proposer, supposer*) et pour leurs constructions, nous rappelons que l'ensemble des verbes du LST est constitué à partir du corpus d'experts Scientext-SHS dans le cadre du projet Termith (Hatier *et al.* 2016). Ce lexique est composé de 342 verbes (*cf.* Chapitre III). Nous avons d'abord observé si les 342 verbes apparaissaient dans les deux corpus : 334 apparaissent dans le corpus d'étudiants chinois contre 336 dans le corpus d'étudiants français. Cette forte présence montre que le LST est central dans ces écrits. On peut en déduire que les apprenants chinois sont conscients d'une variété lexicale des verbes du LST, mais le fait que ces verbes soient présents dans leurs écrits ne veut pas

pour autant dire qu'ils en maîtrisent les constructions et le sens. On se posera ainsi les questions suivantes : quelles sont les constructions verbales mobilisées ? Les verbes sont-ils employés à bon escient au plan sémantique ? Les étudiants perçoivent-ils la polysémie ?

A. Répartition des verbes fréquents du LST dans les corpus Sup-Chinois et Sup-Natifs

Observons maintenant comment se répartissent les 342 verbes du LST. Nous avons projeté cette liste sur le Corpus Sup-Chinois et le Corpus Sup-Natifs et nous avons relevé la fréquence absolue et la fréquence relative (nombre d'occurrences par millions de mots) de ces verbes dans chaque corpus. Ensuite, nous avons calculé le ratio de fréquence relative de ces deux corpus par rapport au Corpus Scientext-SHS (SSHS). À titre d'exemple, nous montrons dans le tableau 5.5 les statistiques obtenues pour le verbe *montrer* :

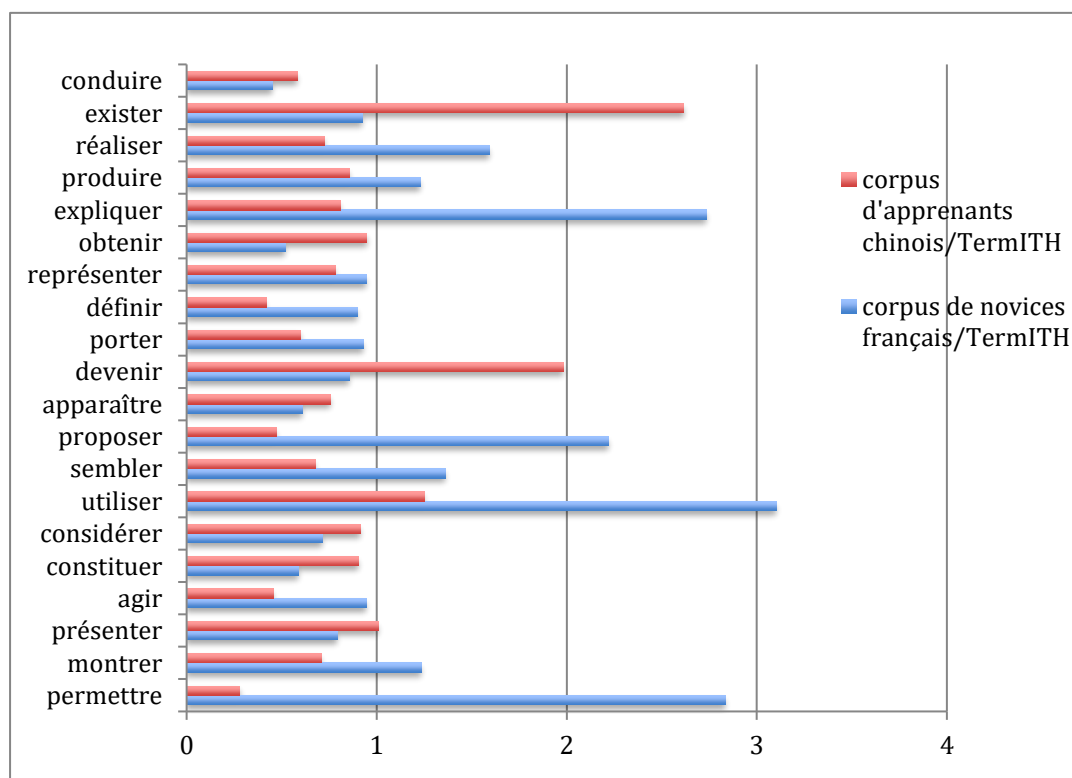
lemme	Fréquence absolue		Fréquence relative		Ratio de fréquence relative	
	Sup-Natifs	Sup-Chinois	Sup-Natifs	Sup-Chinois	Sup-Natifs/SSHS	Sup-Chinois/SSHS
<i>montrer</i>	363	271	780	449	1,235935576	0,711179234

**Tableau 5.5 : Les fréquences absolue/relative et le ratio de fréquence relative pour le verbe *montrer***

Comme on peut le voir, les novices français emploient plus fréquemment le verbe *montrer* (avec une fréquence relative de 780 occurrences par millions de mots) que les apprenants chinois (avec une fréquence relative de 449 occurrences par millions de mots). Si le ratio est  $> 1$ , les verbes sont sur-employés par rapport aux experts ; à l'inverse, ils sont sous-employés. Ainsi, en observant le ratio de fréquence, on constate que les novices français suremployent ce verbe par rapport aux experts (avec un ratio d'environ 1,2), tandis que les apprenants chinois sous-emploient ce verbe (avec un ratio d'environ 0,7). Il nous semble important de rappeler que ce verbe, étudié dans le chapitre IV, est doté d'une valeur démonstrative et métadiscursive, présentant des constructions syntaxiques et des routines verbales variées. Le sous-emploi de ce verbe par les apprenants chinois nous amène à supposer que ses propriétés sémantiques et syntaxiques dans l'écrit scientifique ne sont pas maîtrisées.

Ainsi, les cas de sous-emplois relevés permettent de repérer les verbes à intégrer dans les activités.

Si nous nous intéressons aux 20 verbes les plus fréquents du LST du corpus d'experts, les novices français, et encore plus les apprenants chinois, marquent une tendance au sous-emploi comme l'illustre la figure 5-3 :



**Figure 5-3 : Ratio de fréquence relative pour les 20 verbes les plus fréquents<sup>4</sup> de notre liste du LST**

Si l'on examine ce graphique de plus près, on constate que les ratios de fréquence de 8 verbes se distinguent nettement : *permettre*, *montrer*, *sembler*, *proposer*, *devenir*, *expliquer*, *réaliser*, *exister*. D'une part, les verbes *permettre*, *montrer*, *sembler*, *proposer*, *expliquer* et *réaliser* sont sur-employés par les novices français et sous-employés par les chinois ; d'autre part, les verbes *devenir* et *exister* sont sur-employés par les chinois et sous-employés par les novices français.

Les verbes *permettre*, *montrer*, *proposer*, *expliquer* sont très fréquents et remplissent des fonctions rhétoriques importantes dans l'écrit scientifique (*cf.*

<sup>4</sup> Ces verbes sont classés dans l'ordre croissant selon la fréquence des 20 verbes dans le corpus d'experts.



Chapitre IV). Par exemple, le verbe *permettre* (fréquence relative de 3172 pour le Corpus Sup-Natifs contre 178 pour le Corpus Sup-Chinois) est ainsi largement utilisé dans la construction causative *permettre de*, une construction essentielle dans l'argumentation scientifique (voir les exemples 33 et 34 venant des deux corpus).

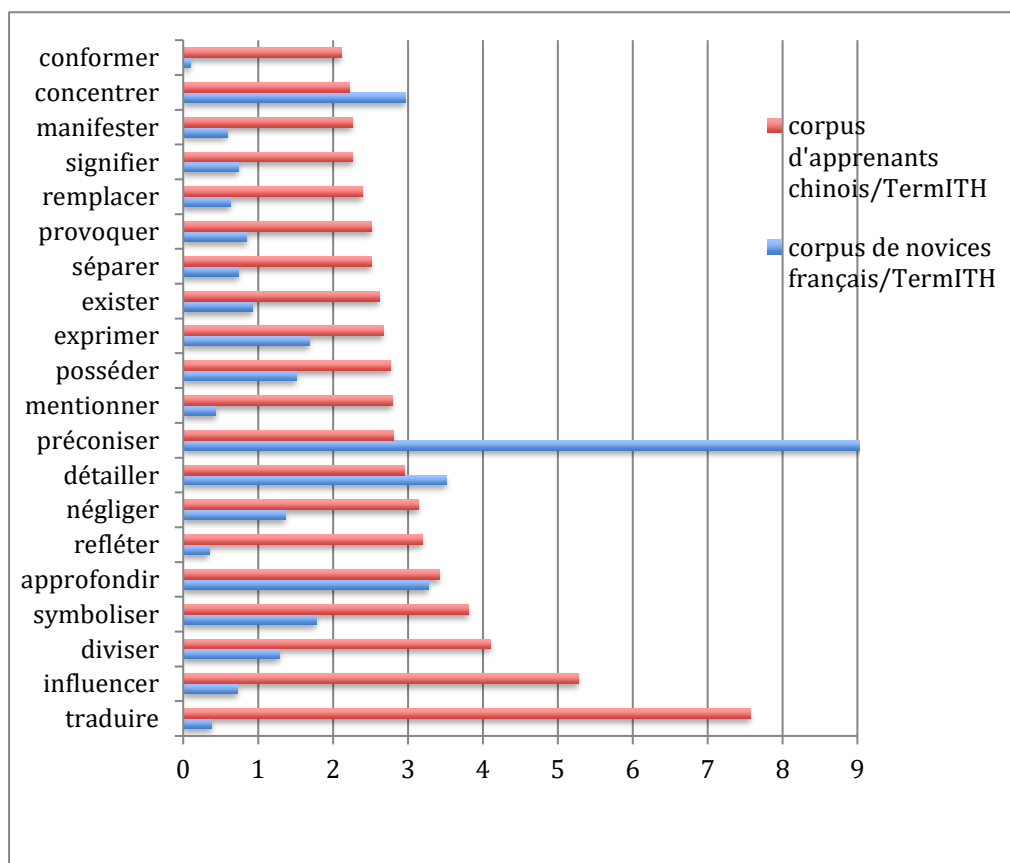
- 33 Alors cette conception **permet d'**expliquer pourquoi l'« aspect aoristique » (la « perfectivité ») se trouve associé aux notions de globalité et de ponctualité. (apprenants chinois\_linguistique\_1)
- 34 De plus, le modèle des rimes **permet de** mettre en évidence l'importance des relations entre les capacités méta phonologiques et l'acquisition de la lecture. (novices français\_6\_BOCH\_M)

Le sous-emploi de ce verbe implique le sous-emploi de la construction causative '*permettre de*', ce qui peut être dû au fait que les apprenants chinois ne maîtrisent pas bien cette construction.

La faible fréquence des verbes *montrer*, *proposer* et *expliquer* dans le corpus Sup-Chinois est peut-être liée au fait que la polysémie de ces verbes n'est pas maîtrisée. Il nous semble important de confirmer notre supposition par l'examen de l'utilisation des constructions verbales associées à ces verbes (voir la section 5.3.3.2.).

#### *B. Les verbes les plus spécifiques du corpus Sup-Chinois*

Dans un deuxième temps, nous souhaitons examiner quels sont les verbes du LST les plus spécifiques par les apprenants chinois. La figure 5-4 donne un aperçu des 20 verbes du LST les plus spécifiques.



**Figure 5-4 : Ratio de fréquence relative pour les 20 verbes du LST les plus utilisés par les apprenants chinois**

De manière générale, à travers l'observation du corpus, on distingue deux cas qui pourraient expliquer le suremploi de ces verbes par les apprenants chinois. D'une part, il peut être lié à la thématique du corpus d'apprenants chinois (notamment, les thèmes de la traduction et de la littérature). Par exemple, *traduire* (fréquence par million de mots : 711 contre 97 pour le corpus de novices français) est le verbe le plus fréquent chez les apprenants chinois. Ceci n'est cependant pas dû à l'emploi de l'acception scientifique (« être l'expression d'un phénomène ») [35], mais à son acception de la langue générale « formuler dans une autre langue étrangère » [36].

35 L'auteur assimile ce principe à une « lecture devinette » qui **se traduit par** une méthode mise en place par l'enfant pour reconnaître à partir d'indices le mot le plus probable. (novices français\_06\_BOCH\_M)

36 Et les traducteurs préfèrent **traduire** ce poème en vers libre au lieu de vers traditionnel et fix. (apprenants chinois\_traduction\_1)

Nous faisons le même constat pour d'autres verbes comme *exprimer*, *refléter*, *symboliser*, *signifier*, *manifester*, *remplacer*. Bien qu'étant des éléments du LST, ils sont souvent employés dans un sens ne s'inscrivant pas dans notre définition du LST, mais qui renvoie à l'interprétation des objets littéraires [37] :

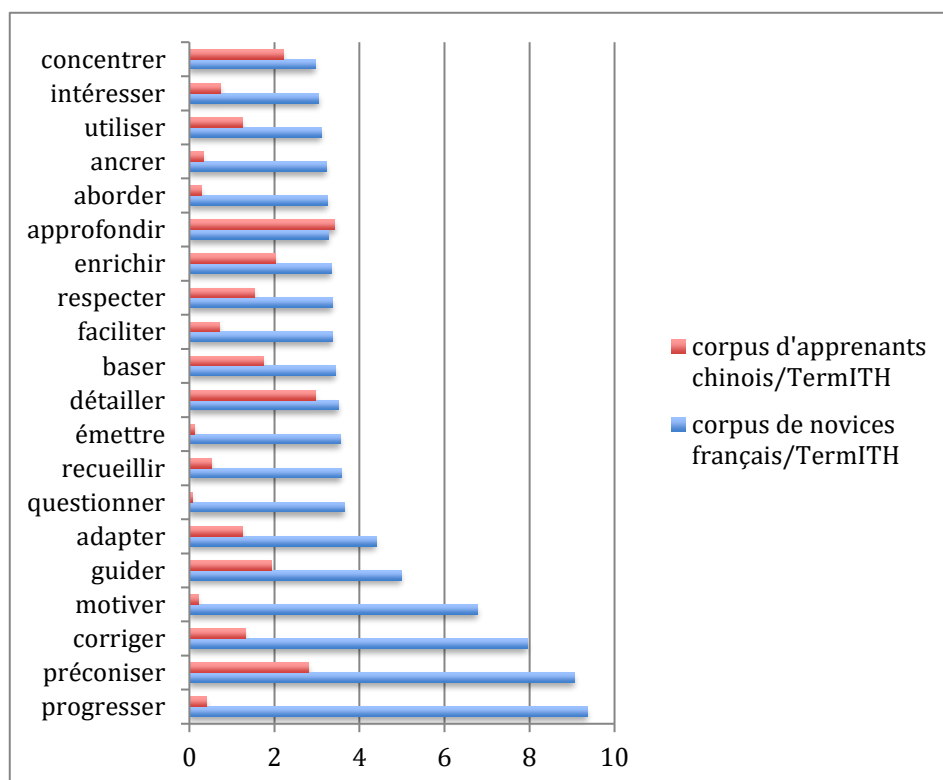
- 37 Pascal essayait d'**exprimer** des préoccupations fondamentales de l'humanité dans cet ouvrage : réflexions sur l'art de persuader, sur la condition misérable de l'homme (apprenants chinois\_littérature\_6)

D'autre part, la sur-utilisation de certains verbes peut également s'expliquer par le répertoire lexical limité des apprenants. Ceci concerne notamment les verbes comme *diviser*, *concentrer*, *mentionner* et *influencer*. Lorsque les apprenants souhaitent exprimer certains sens, ils ont tendance à se cantonner à un même verbe (ou à une même construction verbale). Par exemple, ils utilisent exclusivement l'expression *diviser en* [38] (82/89 occurrences) pour présenter la structure d'un ensemble en plusieurs parties, et l'expression *se concentrer sur* [39] (48/75 occurrences) pour exprimer l'attention qu'ils portent à un objet.

- 38 L'activité esthétique est un processus complexe. Nous la **divisons** simplement **en** trois parties : la forme de langue et l'essence de l'esthétique ; la conscience esthétique du traducteur ; le cours d'esthétique de la traduction (apprenants chinois\_traduction\_5)
- 39 Nous **nous concentrons sur** l'analyse du temps et de l'espace dans La Modification, afin de montrer comment la révolution littéraire de Butor est réalisée. (apprenants chinois\_littérature\_2)

### C. Les verbes les plus spécifiques du corpus Sup-Natifs

Afin de mieux comprendre les différences entre les apprenants chinois et les novices français, nous présentons dans la figure ci-dessous les 20 verbes les plus spécifiques des novices français :



**Figure 5-5 : Ratio de fréquence relative pour les 20 verbes du LST les plus utilisés par les novices français**

Nous avons montré précédemment que certains emplois non-LST de verbes pour lesquels on observe une fréquence élevée dans le corpus d'apprenants chinois sont liés aux thématiques abordées par les apprenants chinois. De la même manière, la thématique du corpus de novices français, la didactique, est fortement représentée parmi les verbes les plus spécifiques : *progresser*, *respecter*, *corriger*, *guider* et *motiver* (voir l'exemple [40]) :

- 40 C'est également grâce à l'enseignant qui, par ses échanges oraux avec les enfants, leurs permet de **progresser** dans la construction vocable et par là même dans l'apprentissage du lexique. (novices français\_13\_BOCH\_M)

L'analyse qualitative des concordances est ici nécessaire pour savoir si les verbes sont effectivement utilisés dans leur sens scientifique ou bien si les novices français et les apprenants chinois ne maîtrisent pas l'emploi de sens spécialisés.

Contrairement au corpus Sup-Chinois, la plupart des verbes surreprésentés dans le corpus de novices français peuvent être regroupés dans nos classes sémantiques de verbes relevant de la description de l'activité scientifique. Ils appartiennent

notamment aux classes sémantiques suivantes : {fondement} (*baser, ancrer*), {analyse} (*approfondir, détailler, aborder*), {collecte des données} (*recueillir*). Il nous semble que les écrits des novices français se rapprochent davantage de ceux des experts sur ce point.

En outre, il est intéressant de constater que les étudiants chinois connaissent tous les verbes, mais ils les utilisent moins que les natifs. Donc, nous pensons qu'il s'agit de lacunes peut-être davantage sémantiques voire donc de lacunes discursives puisque les étudiants chinois ne savent pas quel verbe utiliser à quel moment.

Nous notons ainsi des différences de spécificité qui sont un indice de la variabilité de l'emploi des verbes du LST, et notamment de leur usage spécifique dans l'écrit scientifique. Afin d'étudier ces possibles divergences, nous procédons à une analyse quantitative et qualitative des constructions verbales présentes dans les deux corpus (en confrontant les données aux patrons identifiés dans le corpus d'experts), et ce, pour 10 verbes du LST.

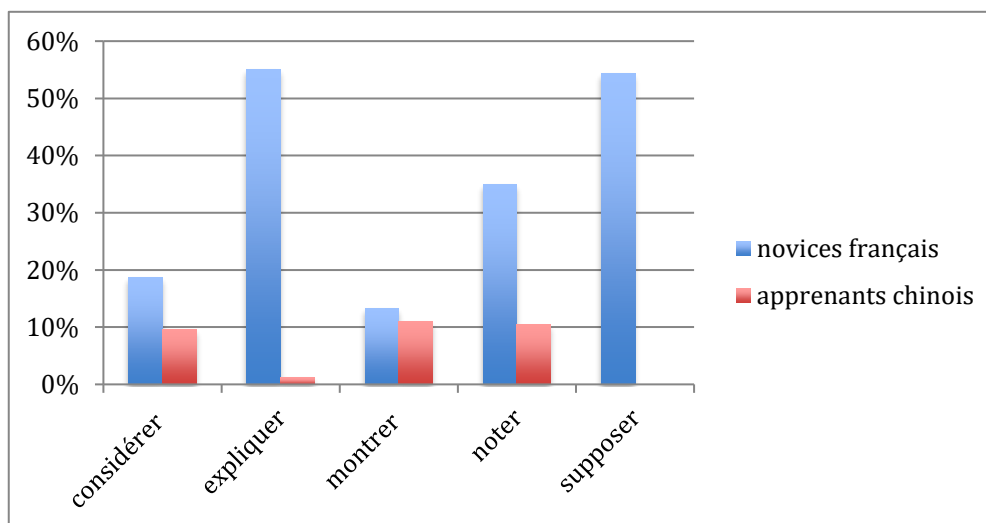
### 5.3.3.2. Fréquences des constructions verbales

Dans cette partie, nous reprenons les 10 verbes retenus pour l'analyse des erreurs et déjà étudiés dans le chapitre IV (*choisir, considérer, comparer, décrire, étudier, expliquer, montrer, noter, proposer, supposer*) afin d'examiner les constructions verbales surreprésentées et sous-représentées chez les apprenants chinois et les novices français. Ces données seront confrontées à nos analyses des patrons verbaux pour mettre en lumière les différences. Nous émettons l'hypothèse que les constructions verbales dans les écrits chinois sont moins variées que celles des novices français et des experts.

#### Construction à sujet humain et en complétive que

Dans un premier temps, nous avons relevé la construction à sujet humain et à complétive *que* dans les deux corpus. Cette construction est mobilisée par cinq des verbes étudiés : (sujet humain) *considère/explique/montre/note/suppose que*. Comme le montre la figure 5-6, les novices français privilégient cette construction de manière assez nette par rapport aux apprenants chinois. Notons par ailleurs que le verbe

*supposer* associé à cette construction est absent de notre corpus d'apprenants chinois.



**Figure 5-6 : La proportion des constructions *Sujet humain + Verbe + que* (*considérer*, etc.)**

Nous avons déjà montré dans le chapitre IV que les verbes associés à cette construction activent des sens spécifiques dans l'écrit scientifique, surtout lorsque le sujet humain renvoie à l'auteur (ou aux auteurs). Au niveau du sémantisme, le sous-emploi de ces verbes dans cette construction implique que les sens associés y sont sous-employés. Par exemple, le verbe *considérer* (*nous considérons que*) prend le sens « se former une opinion », le verbe *noter* (*on note que*) signifie « remarquer, prêter attention à » et le verbe *supposer* veut dire « faire l'hypothèse que ».

Si l'on examine les cinq verbes dans cette construction, on peut constater que les novices français présentent des emplois similaires aux patrons identifiés dans le corpus d'experts. Par exemple, pour le verbe *montrer*, quelques emplois métonymiques au niveau du sujet ont été relevés chez les novices français (9/354 occurrences) (*recherche/étude montre que*), tandis que ces emplois sont quasiment absents (2/272 occurrences) chez les apprenants chinois (voir l'exemple [41]).

- 41 Même si l'école a un rôle majeur dans l'incitation de l'enfant à la lecture, **de nombreuses études ont montré que** l'entourage de l'enfant avait un rôle dans cet apprentissage. (novices français\_04\_BOCH\_M)

Par ailleurs, de même que les experts, les novices français choisissent souvent le pronom *on* ou *nous* comme sujet qui représente l'auteur (ou les auteurs), notamment

pour les verbes *considérer* (11/171 occurrences), *noter* (18/111 occurrences) et *supposer* (30/57 occurrences), (voir l'exemple [42]), alors que les apprenants chinois choisissent beaucoup moins ou très rarement le pronom *on* ou *nous* comme sujet auprès de *noter* (8/96 occurrences), *considérer* (6/290 occurrences), *supposer* (0 occurrence), comme le montre l'exemple [43] dans lequel le sujet renvoie à un autre chercheur.

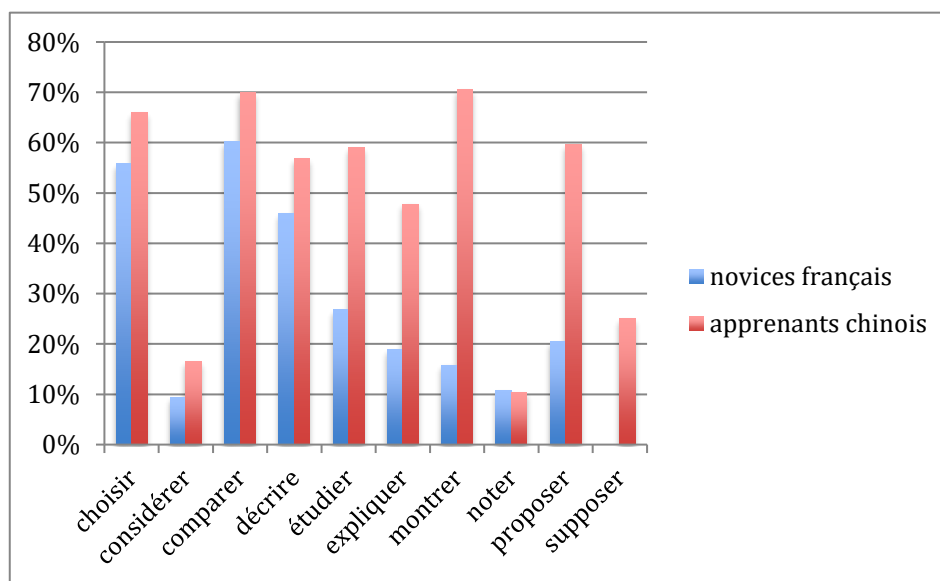
- 42 **D'après l'auteur**, l'apprenti lecteur automatisera progressivement diverses procédures : identification, prise d'indices, anticipation, capacité à faire appel au savoir phonographique, recherche de sens, rappel en mémoire, appel au contexte et synthèse. **On note que** ce sont exactement les procédures développées plus haut par Bellenger, Goigoux et Poslaniec pour l'entrée dans la lecture active. (novices français\_26\_BOCH\_M)
- 43 En ce sens, **Ricci montre que** la Chine est guidée par des philosophes, ce qui contribue à illustrer l'idéal du « roi philosophe ». (apprenants chinois\_socioculturel\_3)

De ce fait, la forme modalisée *on+pouvoir* fréquemment associée aux verbes *considérer*, *supposer* et *noter* dans l'écrit scientifique est quasiment absente chez les apprenants chinois. Il a été démontré que cette forme modalisée permet d'atténuer la prise de position de l'auteur (*cf.* Chapitre IV), ce qui constitue une caractéristique importante de l'écrit scientifique.

Il est intéressant de noter que l'emploi de la construction *on/nous + montrer + que* est faible chez les apprenants chinois (2/272 occurrences) comme chez les novices français (3/354 occurrences). Dans le chapitre IV, nous avons montré que cette construction verbale comporte des valeurs métatextuelle et démonstrative importantes. Il y a donc lieu de penser que ces deux publics ne sont pas conscients de ces valeurs sémantiques.

#### Construction à sujet humain et objet nominal

Contrairement à la construction précédente, la construction à sujet humain et objet nominal relevée pour chacun des dix verbes étudiés semble être privilégiée par les apprenants chinois, comme l'illustre la figure 5-7 :



**Figure 5-7 : La proportion des constructions *sujet humain + choisir + objet nominal (choisir, etc.)***

Notons par ailleurs que pour le verbe *considérer*, la construction retenue est *(sujet humain) considérer (objet nominal) comme*. La principale différence entre les deux publics concerne notamment les verbes *étudier*, *expliquer*, *montrer*, *proposer*, et *supposer*, moins utilisés par les novices français dans ce type de construction.

En observant les lignes de concordances, on remarque que les apprenants chinois ont tendance à choisir un nom non-LST comme objet et que les verbes sont souvent utilisés dans un sens appartenant à la langue générale. Comparons les exemples [44] et [45] :

- 44 La première catégorie : des poèmes concernent l'enfance et l'adolescence de Léopoldine où **le poète décrit la naïveté, la tendresse et la sagesse** de sa fille. (apprenants chinois\_literature\_10)
- 45 **Il décrit la méthode** qu'il a choisie comme « un apprentissage direct intentionnel utilisant visualisation et copie ». (novices français\_41\_BOCH\_M)

Les noms *naïveté*, *tendresse*, *sagesse* n'appartiennent pas au LST tandis que le nom *méthode* fait partie de moyens mis en place pour effectuer une recherche scientifique. Par rapport aux apprenants chinois, les novices français emploient davantage des noms objets relevant du LST, se rapprochant ainsi des experts. Prenons l'exemple du verbe *montrer*, utilisé par les experts pour des objets-noms renvoyant à



une qualité (*importance, rôle*) ou une relation (*lien, différence*) (cf. Chapitre IV). Ces emplois ont également été relevés chez les novices français (6/354 occurrences) mais sont quasiment absents chez les apprenants chinois. Comparons les exemples [46] et [47] :

- 46 M. Métra en 2010 **montre le rôle** important de l'enseignant en tant qu'instaurateur d'un cadre sécurisant. (novices français\_16\_BOCH\_M)
- 47 Dans lequel, **l'auteur nous montre la vie** des différents personnages qui n'ont ni de noms ni de caractères personnels (apprenants chinois\_literature\_5)

Dans l'exemple [47], le verbe *montrer* est utilisé dans le sens de 'faire imaginer, représenter à l'esprit' et non pas dans le sens de 'démontrer'. Ces différences d'emploi peuvent s'expliquer par plusieurs facteurs : d'abord, comme nous l'avons déjà mentionné, les thématiques des mémoires d'apprenants chinois sont différentes de celles des novices français. Ensuite, les écrits des apprenants chinois relèvent plutôt du type descriptif que d'un type argumentatif. Or, il s'agit de mémoires de fin d'études de master, qui devraient bien au contraire être du type argumentatif. C'est pourquoi nous considérons que ces mémoires ne sont pas de véritables travaux de recherche. Enfin, il apparaît que les apprenants chinois ne sont pas familiarisés avec les sens du verbe appartenant à l'écrit scientifique, ignorant ainsi les rôles de ces verbes dans ce type d'écrit.

Par ailleurs, rappelons que dans la section 5.2.3., nous avons relevé des erreurs de collocations verbales chez les apprenants chinois. Les apprenants chinois sont confrontés à davantage de difficultés que les novices français concernant le choix des objets nominaux (voir les exemples [48] et [49]) :

- 48 Premièrement, la traduction interlinguale ou la reformulation, c'est-à-dire traduire une langue à l'aide de la même langue, généralement quand **on explique une chose** dans son propre pays. (apprenants chinois\_traduction\_3) (Correction proposée : quand on explique un terme spécialisé ou compliqué)
- 49 A ce cause, **on propose quelques tactiques** accessibles en vue de donner une traduction fidèle et de bonne qualité dans ce mémoire. (apprenants chinois\_traduction\_4) (Correction proposée: Ainsi, dans ce mémoire, on propose des

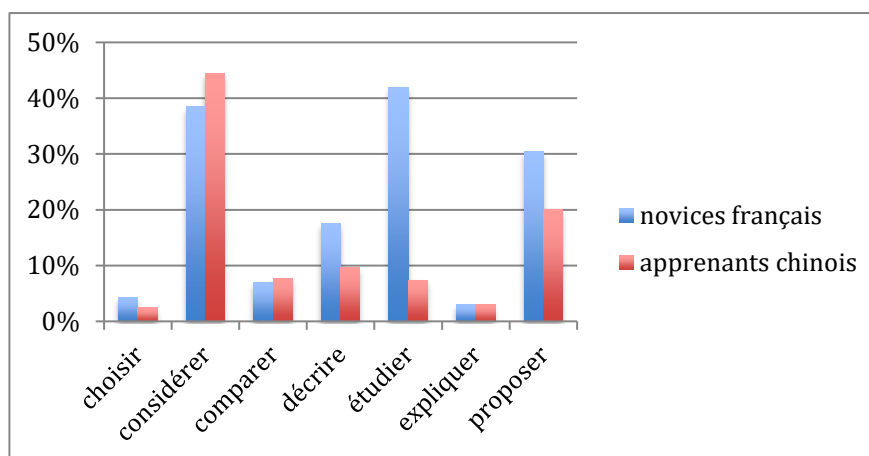
méthodes de traduction permettant de restituer une traduction fidèle et de bonne qualité.)

Dans l'exemple [48], l'objet nominal *une chose* est un terme imprécis alors que dans l'exemple [49], l'emploi du mot *tactique* est maladroit et ne convient pas dans ce contexte.

En bref, les apprenants chinois emploient plus fréquemment la construction à sujet humain et objet nominal que les novices français. Ils se distinguent des novices français et des experts par le fait qu'ils choisissent souvent des objets noms non-LST et emploient le verbe dans un sens éloigné de l'écrit scientifique, ce qui nous conduit à penser que même s'ils connaissent les verbes étudiés, ils ne seraient pas pour autant conscients de leurs sens et de leurs collocations dans le contexte scientifique, car cette dimension ne leur est pas enseignée.

#### Construction au passif

Il nous semble important de rappeler que dans le chapitre IV, nous avons montré que ces verbes fréquemment employés au passif ou au passif réduit (*modèle décrit/étudié/proposé*) permettent de mettre en avant l'objet à analyser : *les résultats sont décrits dans la section suivante, notre choix peut être expliqué par le fait que, l'exemple étudié nous permet de dire que, l'approche proposée s'inscrit dans, etc.* Nous avons également expliqué dans le chapitre III que les formes passives dans l'écrit scientifique ont pour fonctions principales d'« assurer une progression à thème constant », d'« effacer le sujet, alléger la structure phrastique » et de contribuer à « l'impersonnalité du texte technoscientifique » (Riegel et al. 2009 ; Kocourek 1991 ; cité par Fifielska 2015, p. 58). Afin de savoir si la construction au passif (y compris le passif réduit) est maîtrisée par les apprenants chinois, nous avons examiné cette construction dans les deux corpus, à travers sept verbes indiqués dans la figure 5-8.



**Figure 5-8 : La proportion des constructions passives**

On remarque que les 3 verbes *décrire*, *étudier* et *proposer* sont davantage utilisés au passif (ou au passif réduit) par les novices français. Notons par ailleurs que le passif réduit est beaucoup moins fréquent chez les apprenants chinois : on relève pour le verbe *décrire* 10/237 occurrences chez les apprenants chinois contre 12/102 occurrences chez les novices français, pour le verbe *étudier* 3/135 occurrences contre 29/119 occurrences et pour *proposer*, 13/129 occurrences contre 102/473 occurrences.

Comparons les deux exemples ci-dessous :

- 50 La répartition des erreurs selon **la typologie décrite ci-dessus**, montre que l'orthographe grammaticale a représenté la principale source de difficultés des étudiants en situation de production écrite, ce qui témoigne de besoins importants dans ce domaine. (novices français\_41\_BOCH\_M)
- 51 Même au cours de l'écriture, l'auteur ne possède pas l'autorité absolue sur son œuvre. Il écrit le texte, le texte a aussi de l'influence sur lui, ils sont situés dans un cycle d'interaction, **nous pouvons décrire leur relation** par le mot intersubjectivité. (apprenants chinois\_traduction\_8) (Correction proposée : Pendant la rédaction....le texte exerce aussi une influence sur l'auteur, si bien qu'on pourrait dire qu'ils sont en interaction l'un avec l'autre. Cette relation peut être considérée comme intersubjective)

Dans l'exemple [50], le passif renvoie à la description précédente, permettant de rendre la phrase moins lourde. Dans l'exemple [51], on comprend bien que l'apprenant mentionne la relation entre l'auteur et son texte, mais celle-ci n'est pas le

sujet dans la phrase *nous pouvons décrire leur relation*. Il semble qu'ici, la forme passive conviendrait mieux puisque, d'une part, elle enlèverait le caractère personnel donné par le *nous* à la forme active, atténuant ainsi l'implication de l'auteur en tant que sujet ; d'autre part, elle maintiendrait la cohésion textuelle, permettant de centrer le propos sur la relation entre l'auteur et son texte. D'ailleurs, l'exemple [51] présente des erreurs sémantiques liées au verbe.

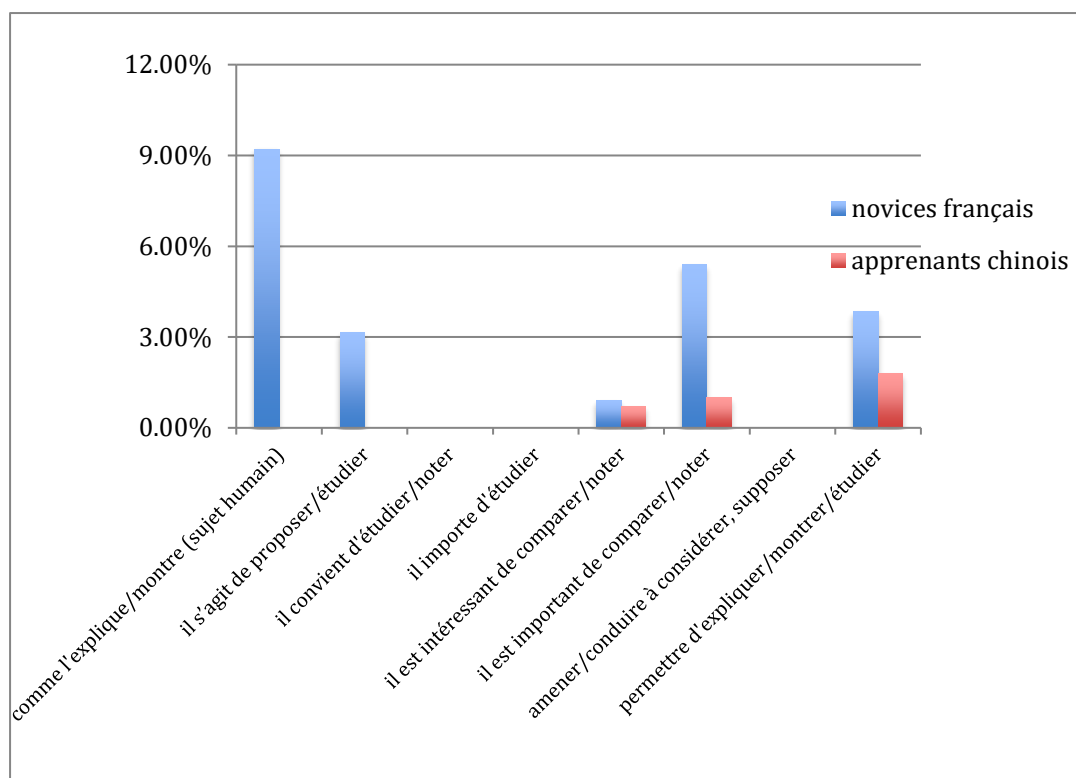
En revanche, on observe que le verbe *considérer* utilisé dans la construction *être considéré comme* semble être bien maîtrisée par les apprenants chinois [52] :

- 52 L'imparfait textuel **est considéré comme** l'arrière-plan opposant au premier-plan (passé simple), par la description de l'imparfait dans notre manuel pédagogique, ce courant est simple à comprendre, il ne sera pas analysé. (apprenants chinois\_linguistique\_1)

Il est important de souligner que ce verbe est classé comme un verbe d'opinion, parmi les verbes étudiés dans le chapitre IV. Les verbes souvent utilisés au passif sont des verbes d'analyse dont les objets sont fortement thématiques. Ainsi, il nous semble que cette tendance de thématisation n'est pas prise en compte par les apprenants chinois.

#### Constructions sous-employées ou absentes dans le corpus d'apprenants chinois

En dehors des constructions verbales principales présentées ci-dessus, nous avons également examiné, pour les dix verbes retenus, certaines constructions spécifiques à l'écrit scientifique relevées dans le corpus d'experts dans le chapitre IV. Il s'agit de trois types de constructions : les constructions en *comme* (*comme l'explique/le montre X*) ; les contextes impersonnels du type *il s'agit/convient/importe de* (*proposer, étudier, noter*) et du type *il est intéressant/important de* (*comparer, noter*) ; les constructions causatives *amener/conduire à* (*considérer, supposer*) et *permettre de* (*montrer, étudier, expliquer*). Nous souhaitons donc observer la fréquence et la diversité de ces constructions dans le corpus Sup-Chinois et le corpus Sup-Natifs. Les résultats apparaissent dans la figure 5-9 :



**Figure 5-9 : La proportion des constructions sous-employées ou absentes chez les apprenants chinois**

De manière générale, deux cas peuvent être distingués : d'une part, les contextes impersonnels *il est important/intéressant de (comparer, noter)* et les constructions causatives *permettre d'expliquer/montre/étudier* sont sous-employées par les apprenants chinois. Comme expliqué dans le chapitre IV, les premières expriment la nécessité ou l'intérêt lié à la comparaison des données et à la prise de conscience d'une information supplémentaire alors que les secondes servent à mettre en évidence le processus de raisonnement et à justifier le choix et l'analyse de l'auteur. Le sous-emploi de ces constructions nous conduit à supposer leur non-maîtrise ou tout au moins une maîtrise insuffisante de la part des apprenants chinois. L'exemple suivant semble confirmer notre hypothèse :

- 53 **Voir l'histoire de l'auto-traduction, il est important de noter que** le français persistait parmi les élites, par conséquent, l'activité de la traduction entre le français et l'allemand, l'anglais, l'italien, le portugais et d'autres langues ont été largement exercée pendant une longue période. (apprenants chinois\_linguistique\_1) (Correction proposée : Lorsqu'on regarde l'histoire de l'auto-traduction, on peut constater que le français était la langue des élites. Par conséquent, l'activité de la traduction entre le

---

français et d'autres langues a été largement pratiquée pendant une longue période.)

Dans l'exemple [53], la construction *il est important de noter que* ne convient pas dans ce contexte puisque celle-ci est employée dans l'écrit scientifique pour inviter le lecteur à porter son attention sur une information supplémentaire, tandis qu'ici, il s'agit d'avantage d'un constat général.

D'autre part, les constructions en *comme*, les contextes impersonnels du type *il s'agit/convient/importe de (proposer, étudier, noter)* et les constructions causatives *amener/conduire à (considérer, supposer)* sont totalement absentes chez les apprenants chinois. Ces constructions jouent des rôles rhétoriques importants dans l'argumentation et le raisonnement de l'écrit scientifique (*cf.* Chapitre IV). Par exemple, les constructions en *comme* permettent de renforcer l'argumentation (*comme l'explique le chercheur X*) ou de guider le lecteur dans la lecture (*comme nous l'avons montré plus haut*). En outre, il est intéressant de noter que les constructions *il convient/importe de (proposer, étudier, noter)* et *amener/conduire à (considérer, supposer)* sont également absentes chez les novices français. Cette absence peut être liée à la taille des deux corpus, mais peut aussi révéler que les étudiants français comme les apprenants chinois ont une connaissance limitée de l'emploi de ces constructions verbales. On peut avancer que ces deux publics sont confrontés à fois à la diversité et à la complexité des constructions verbales en tant que novices dans ce genre discursif. Autrement dit, ils n'ont pas à leur disposition les (ou pas suffisamment de) moyens linguistiques permettant d'aborder le sujet, de justifier le choix, de renforcer l'argumentation, etc. Ces lacunes peuvent également être liées au fait qu'ils n'ont pas encore construit une représentation précise du fonctionnement de la langue scientifique.

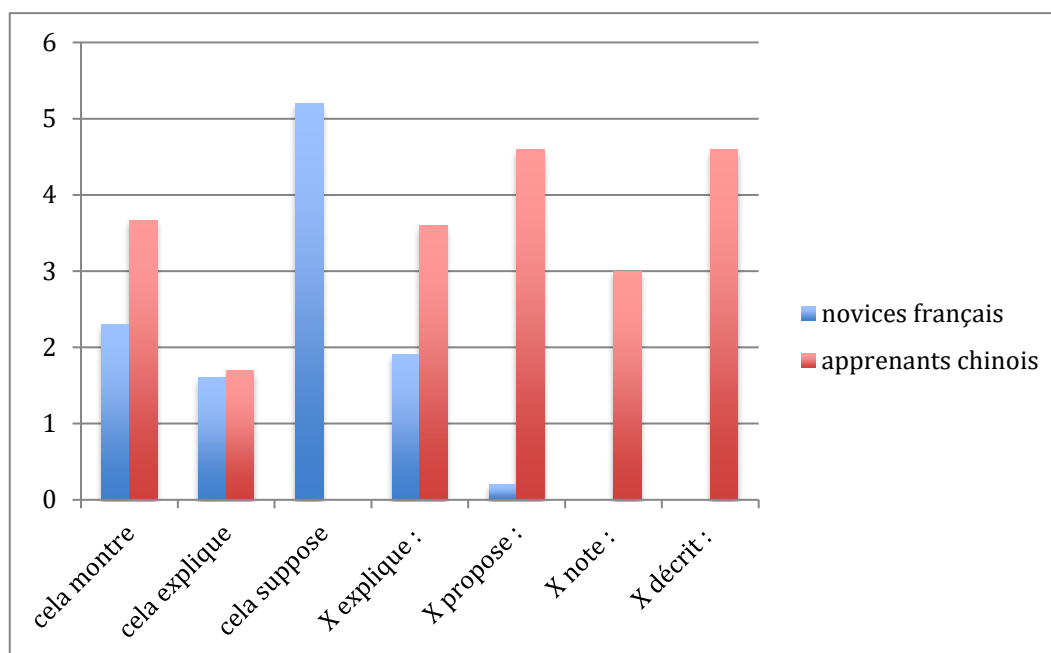
Pourtant, les résultats nous montrent que les écrits des novices français se rapprochent davantage de ceux des experts, tant du point de vue quantitatif que qualitatif. Une meilleure maîtrise de la langue facilite sans conteste leur rédaction scientifique. Ce n'est pas le cas des étudiants chinois qui n'ont pas appris à argumenter dans le système éducatif chinois et sont moins sensibilisés au genre textuel d'écrit scientifique. Notre diagnostic rejoint l'étude de Bi (2016) qui a mis en évidence des écarts rhétoriques entre les pratiques des écrits argumentatifs en français

et en chinois en mettant en évidence trois niveaux de difficultés des étudiants chinois : les connaissances linguistiques, l'organisation textuelle mais aussi les stratégies argumentatives.

Pour conclure, on relève des difficultés pour les deux publics dans la maîtrise de la diversité des constructions verbales, qui sont nettement accentuées pour les apprenants chinois. En effet, l'état d'interlangue très imparfaite, associé à la non-maîtrise du genre de l'écrit scientifique et au déficit de la compétence argumentative, constitue un handicap qui gêne considérablement les apprenants chinois dans la rédaction académique.

#### *Constructions surreprésentées dans le corpus d'apprenants chinois*

Lorsqu'on observe la fréquence et la diversité des constructions verbales récurrentes pour les dix verbes retenus respectivement dans les corpus Sup-Chinois et Sup-Natifs, on constate que les apprenants chinois ont tendance à suremployer certaines constructions, ce qui est révélateur des défauts des novices de l'écrit scientifique. Il s'agit principalement de deux types de constructions : la construction verbale ayant le pronom démonstratif *cela* comme sujet (*cela montre que*) et la construction dans laquelle le verbe est suivi d'une citation directe (*l'auteur explique : « .... »*). La figure 5-10 présente un aperçu de chacune de ces constructions dans les deux corpus. De manière générale, le premier type de construction est employé par les deux publics tandis que le deuxième apparaît plus spécifique aux apprenants chinois.



**Figure 5-10 : La proportion des constructions suremployées chez les apprenants chinois**

La construction verbale ayant le pronom démonstratif *cela* en sujet est notamment associée aux verbes *montrer* et *expliquer* chez les apprenants, comme l'illustrent les exemples [54] et [55].

54 En outre, les écrivains étrangers sont devenus les écrivains du monde après que ses œuvres avaient été publiés en France. **Cela montre bien la principal raison** pour la prospérité de la culture et la littérature française depuis une longue période. (apprenants chinois\_traduction\_6) (Correction proposée : ce phénomène constitue une raison principale de la prospérité...)

55 Comment s'intégrer dans la culture et la vie du pays immigré, comment être admis par les habitants locaux, **tout cela montre les problèmes** de l'intégration sur l'immigration. (apprenants chinois\_socioculturel\_4) (Correction proposée : toutes ces questions constituent des problèmes de l'intégration des immigrés)

Cependant, dans les deux exemples ci-dessous, l'emploi du pronom démonstratif peut provoquer certaines ambiguïtés référentielles.

Chez les novices français, la construction avec le sujet *cela* est mobilisée par trois verbes *montrer*, *expliquer* et *supposer* (voir l'exemple [56]).



- 56 La définition d'être attentif c'est selon le psychologue W. James (1983) ; cité par Mialet (1999), et repris par Le Manchec (2005) « prendre possession par l'esprit d'un objet ou courant de pensée... ». **Le Manchec (2005) explique que cela suppose** trois grands types d'opérations : de sélection, de tri et d'abandon d'autre chose. (novices français\_12\_BOCH\_M)

Notre constat rejoint l'étude de Boch et Rinck (2016) qui s'intéresse à l'usage des anaphores démonstratives (*ceci, cela, celui-ci, ce N, etc.*) dans un sous-corpus de « Littéracie avancée » (corpus de novices français) en le comparant aux pratiques d'experts. Les auteurs ont remarqué que le pronom *cela* est surreprésenté chez les étudiants français par rapport aux experts. Les premiers préfèrent ne pas spécifier l'anaphore tandis que les experts emploient fréquemment les « anaphores résomptives » du type *Ces analyses permettent* qui mettent en valeur la fonction de synthèse conceptuelle et qui contribuent à l'argumentation dans l'écrit scientifique.

De notre point de vue, sur le plan lexical, l'emploi du pronom démonstratif *cela* peut être considéré comme un moyen détourné de pallier le manque de vocabulaire et de collocations verbales comportant une anaphore nominale (*Cette hypothèse peut, Cet aperçu global indique, Ce point de vue rejoint, etc.*). Sur le plan discursif, comme le soulignent Boch et Rinck (*ibid.*), l'emploi du *cela* révèle que les anaphores résomptives sont peu maîtrisées par les étudiants et que les étudiants n'ont pas encore acquis la compétence à reformuler, à résumer une portion du discours.

Considérons maintenant la construction dans laquelle le verbe est suivi d'une citation directe. Cette construction surreprésentée chez les apprenants chinois montre clairement que ceux-ci préfèrent recourir à des citations directes dans leurs écrits, comme en témoignent les exemples [57] et [58].

- 57 **Molendijk propose** : « l'antécédent temporel d'une phrase P est un fait auquel P est relié à l'aide d'un rapport textuel » [...] (Arie Molendijk, Cahiers Chronos 1 – Anaphores temporelles et (in-)cohérence, p.113) (apprenants chinois\_linguistique\_1)
- 58 Il a ajouté que : la France a appuyé la création d'une force multinationale en Afrique, pour les africains de résoudre leurs propres choses. En même temps, **il a expliqué que** : « N'importe comment réduire le nombre de troupes stationnées en France, la

coopération militaire française et africaine va se poursuivre » 22. (apprenants chinois\_socioculturel\_1)

Dans ces deux exemples, *Molendijk propose* et *il a expliqué* introduisent des citations directes et longues. Ici, les verbes *proposer* et *expliquer* ont introduit le discours d'autrui d'une manière très maladroite et les apprenants ont évité d'exprimer leur propre point de vue, alors que les citations auraient très bien pu être employées pour appuyer leur argumentation.

Sur le plan lexical, la surreprésentation de ces constructions semble rejoindre notre analyse précédente qui met en évidence la sous-représentation des constructions en *comme* (*comme l'explique/le montre le chercheur X*) permettant d'intégrer le discours d'autrui dans son texte. Sur le plan discursif, l'emploi de ces constructions illustre, d'une part, la difficulté des apprenants chinois en matière de représentation de soi en tant qu'auteur ; d'autre part, la non-maîtrise de la reformulation du discours d'autrui pour en faire un argument d'autorité (Boch, 2013).

Dans le corpus d'apprenants chinois, on rencontre fréquemment des citations directes massives où l'apprenant fait une accumulation de citations sans commentaires personnels. Comme le font remarquer Boch et Grossmann (2002), cette pratique démontre que l'auteur est incapable de dominer les différents points de vue pour illustrer ou appuyer une affirmation, traduisant également « un sentiment d'insécurité » (voir l'exemple [59]).

- 59 **Touratier considère que** « l'unité morphologique d'imparfait [...] doit être associée à un signifié non proprement temporel comme "non- actuel" plutôt qu'au signifié "passé" ». **Pour Le Goffic**, « l'imparfait est le temps de l'actualité transposée dans un monde hors d'atteinte, temps de l'inaccessible et non du passé en tant que tel ». **Coseriu observe également que** « dans le système verbal des langues romanes, [...] l'imparfait n'est pas à proprement parler un temps du passé mais uniquement le centre (et le terme neutre) du plan inactuel et [...] s'oppose tout d'abord à l'ensemble du plan actuel, dont le centre est le présent ». (apprenants chinois\_linguistique\_1)

Il nous paraît intéressant de noter que la construction dans laquelle le verbe est suivi d'une citation directe est également relevée chez les étudiants français,

notamment pour le verbe *expliquer*, comme l'illustre l'exemple [60] :

- 60 De même, Garcia-Debanc (1988 : 12) **explique** : « Les enfants se lancent directement dans l'écriture, sans envisager l'organisation d'ensemble de leur texte ou en expliciter les enjeux. Ils rédigent en ajoutant chaque phrase à la phrase immédiatement précédente, dans une avancée pas à pas ». (novices français\_39\_BOCH\_M)

Plusieurs chercheurs ont souligné l'usage abondant des discours directs cités dans les productions des étudiants natifs, qui sont parfois mal compris, voire inutiles (Pollet & Piette, 2002). Selon Boch et Grossmann (2002), la présence forte de la citation est problématique au niveau de l'argumentation et est caractéristique des écrits des novices, contrairement à ceux des experts qui préfèrent l'évocation ou la reformulation de la source. À ce sujet, Mroué (2014) explique que l'usage abondant des citations est lié d'une part, au « stéréotype répandu selon lequel les citations témoignent d'une bonne documentation » (*ibid.*, p. 54), et d'autre part au fait que « citer peut servir de stratégie d'évitement du plagiat » (*ibid.*). Ainsi, on peut penser que les deux publics sont confrontés à la même difficulté d'intégrer le discours d'autrui dans leur texte, mais il apparaît que les maladresses chez les apprenants sont plus importantes du fait que ces derniers emploient souvent une construction maladroite, introduisant la citation de manière massive et abrupte (comme dans l'exemple [59]).

#### 5.4. Récapitulatif du diagnostic

À l'issue de l'analyse des spécificités, nous observons que nous avons obtenu des résultats comparables à ceux de l'étude de Granger et de Paquot (2009b). Nous avons mis en évidence que le statut de novice en matière de rédaction scientifique confronte de manière identique les étudiants natifs et non-natifs aux difficultés de l'usage des verbes du LST et de leurs constructions. Les étudiants natifs connaissent peut-être mieux la langue, mais ne maîtrisent pas encore les normes du genre. Les apprenants chinois produisent des spécificités davantage du côté du sous-emploi des verbes du LST et des constructions. Cela semble en corrélation avec le développement en cours de leur interlangue, contrairement aux étudiants natifs.

Comme Gilquin et Paquot (2008), nous pensons que les écrits académiques des étudiants natifs occupent une place intermédiaire entre ceux des apprenants chinois et ceux des experts.

Si l'on considère les erreurs analysées, on se rend compte que les difficultés sont plus grandes pour les apprenants chinois, du fait de leur statut d'apprenants du français, qu'il s'agisse des confusions de registre, des erreurs de collocations ou des expressions/tourneurs maladroites. Nous rejoignons l'idée de Granger et Paquot (2009) selon laquelle les difficultés des étudiants non-natifs méritent d'être prises en compte et traitées spécifiquement puisqu'ils ont des problèmes liés à la maîtrise de la langue.

Nous proposons ci-dessous un tableau récapitulatif des résultats de diagnostic.

Diagnostic	Cas
Analyse des erreurs	Les erreurs relevées sont relatives aux propriétés sémantique et syntaxique des verbes avec un total de 83% parmi lesquelles la structure syntaxique erronée représente 34% des cas
Observation des cas de suremploi	<ol style="list-style-type: none"> <li>1. Verbes suremployés : <i>devenir, exister, diviser</i>, etc.</li> <li>2. Constructions limitées : <i>diviser en, se concentrer sur</i>, etc.</li> <li>3. Constructions suremployées : construction à la forme active avec sujet humain et objet non animé, <i>cela montre que, il souligne/propose</i> : « ... »</li> </ol>
Observation des cas de sous-emploi	<ol style="list-style-type: none"> <li>1. Verbes sous-employés : <i>montrer, permettre, proposer, expliquer</i>, etc.</li> <li>2. Constructions sous-employées ou absentes : construction en complétive <i>que</i> avec comme sujet les pronoms <i>on</i> ou <i>nous</i>, passif, passif pronominal, construction comportant <i>comme</i> (<i>comme on le montre</i>), <i>permettre</i> (<i>permet d'expliquer</i>), constructions impersonnelles ou contextes impersonnels (<i>il convient de noter que</i>)</li> <li>3. Fonctions rhétoriques sous-employées ou absentes : guider le lecteur (<i>comme nous l'avons montré/expliqué/souligné plus haut</i>) ; renvoyer à d'autres travaux (<i>comme le souligne X</i>) ; exprimer la nécessité ou l'intérêt de l'activité scientifique, (<i>il importe de proposer/étudier, il est intéressant de comparer/constater</i>) ; apporter une justification ou argument (<i>permettre d'expliquer, amener/conduire à supposer</i>) ; exprimer son point de vue, reformuler le discours d'autrui pour illustrer et appuyer une affirmation</li> </ol>

**Tableau 5.6 : Récapitulatif de diagnostic**

Il est important de souligner que ces résultats de diagnostic ont été corroborés par d'autres expérimentations que nous avons menées pendant un cours de Français sur Objectifs Universitaires (désormais FOU) dispensé auprès de six étudiants chinois de Master 2 en SHS à l'Université Grenoble Alpes (voir la section 5.5). Le relevé des erreurs et des cas de sous-emploi effectué apparaît tout à fait comparable concernant

l'utilisation des constructions verbales.

Rappelons que les analyses des erreurs et des spécificités ont pour objectif de mieux outiller l'E/A des constructions verbales du LST. Nous proposerons dans la section suivante, des pistes didactiques.

## 5.5. Pistes didactiques

Afin de remédier aux lacunes relevées, nous envisageons de proposer deux types de ressources qui pourraient aider les apprenants chinois dans leur rédaction académique.

- Une ressource lexicale, inspirée du modèle *Corpus Pattern Analysis*, et intégrant des informations sur les collocations, constructions et routines verbales (voir les tableaux des patrons lexico-syntaxiques de 32 verbes fréquents et les figures des routines de ces verbes avec une dizaine de fonctions rhétoriques dans le chapitre IV)
- Des activités d'entraînement basées sur le corpus pour les apprenants : activités en présentiel guidées par l'enseignant que nous allons présenter dans les parties 5.5.2. et 5.5.3.

Le premier type de ressource est constitué d'une base de données que les enseignants de FLE pourraient utiliser pour l'enseignement du lexique de l'écrit scientifique. Le second peut être considéré comme un exemple illustrant comment l'utilisation des données des corpus en classe de langue peut contribuer à l'enseignement des constructions verbales du LST.

Dans le chapitre II, nous avons expliqué que le corpus Scientext-SHS peut être utilisé en amont dans la description linguistique pour la classe, et en aval dans des exploitations didactiques concrètes. Dans les parties 5.5.2. et 5.5.3., en nous basant sur notre ressource lexicale des patrons verbaux ainsi que sur les résultats de diagnostic, nous souhaitons proposer quelques pistes d'exploitations didactiques concrètes en insistant sur l'importance de l'utilisation des corpus en classe de langue.

Dans un premier temps, nous aborderons les trois façons d'enseigner les corpus. Ensuite, nous présenterons deux expérimentations effectuées avec des mastérisants chinois, qui nous ont permis de prendre du recul et de proposer des pistes didactiques mieux adaptées.

### 5.5.1. Enseignement des patrons avec le corpus

À l'instar de Chambers (2010) et Boulton (2011), nous mettons l'accent dans ce travail sur l'intérêt de l'utilisation du corpus en classe de langue pour enseigner un lexique particulier. L'utilisation du corpus n'est pas toute récente, Johns (1991) en est considéré comme le pionnier à travers son *Data-Driven Learning (DDL)* ou apprentissage guidé par les données. Selon McEnery et Xiao (2010), le corpus donne accès à une « approche opérationnelle des trois 'i' », c'est-à-dire, « Illustration – Interaction – Induction » :

Par illustration on entend le fait de présenter à la classe d'étudiants en langues étrangères, des données réelles ; l'interaction concerne la deuxième phase du travail des étudiants qui échangent leurs opinions et leurs observations ; enfin l'induction permet à l'étudiant de se créer sa propre « règle » (qui sera affinée au fur et à mesure que le corpus de données s'élargit) sur le phénomène étudié. (Di Vito, 2013, p. 160)

Il nous semble donc important de mettre en avant ces approches opérationnelles dans l'enseignement des phénomènes phraséologiques. Nous allons revenir sur cet aspect dans les parties 5.5.2 et 5.5.3. Dans un contexte d'enseignement de langue spécialisée, telle que la langue de l'écrit académique, les chercheurs s'accordent depuis longtemps déjà sur le choix de documents pédagogiques spécialisés ; il en va de même pour les corpus (Boulton & Landure, 2016). Il est préférable d'utiliser un corpus qui représente la spécialité du domaine de l'apprenant pour l'E/A de la langue spécifique à ce domaine. Dans le domaine de l'écrit académique, de nombreuses études empiriques (Yoon & Alain, 2004 ; O'Sullivan & Chambers, 2006 ; Sun, 2007 ; O'Sullivan, 2010 ; Granger et Paquot, 2015) ont montré que l'utilisation des corpus est bénéfique à l'E/A de l'écrit que ce soit de façon directe ou indirecte.

Néanmoins, la pratique de l'enseignement des langues étrangères n'a pas encore

beaucoup utilisé les corpus, notamment dans le domaine du FLE. Comme souligné par Ciekanski (2014, p. 112) : on constate « un écart entre le consensus théorique qui se dégage des recherches menées sur l'utilisation effective des corpus et leur intégration dans les pratiques enseignantes en contexte d'apprentissage », ce qui s'explique par plusieurs éléments. Du côté de l'enseignant, il est nécessaire de former les enseignants à la méthodologie de linguistique de corpus tant sur les plans théorique, que méthodologique et technique (Kübler, 2014). Par ailleurs, l'utilisation des corpus en cours de langue demande à l'enseignant d'assurer un rôle de guide afin d'aider l'apprenant à avancer dans son apprentissage. Comme le montrent plusieurs chercheurs (Di Vito, 2013 ; Giuliani & Hannachi, 2010), l'apprentissage insuffisamment guidé risque de désorienter et de décourager l'apprenant, notamment lorsqu'il s'agit d'éléments linguistiques complexes à observer et à analyser.

Du côté de l'apprenant, la question du développement de l'autonomie se pose. Autrement dit, l'apprenant doit apprendre à répondre à trois questions générales : Pourquoi dois-je utiliser les corpus ? Que dois-je chercher et observer dans les corpus ? Comment chercher et interpréter les données des corpus ? On pourrait ajouter à ces questions le manque de modèles de description linguistique et de corpus pour le français. Parmi ces éléments, l'enseignant joue un rôle important dans l'intégration des corpus en classe, non seulement en termes de préparation du matériel pédagogique à partir des corpus mais aussi dans l'accompagnement de l'apprenant vers l'utilisation et peut-être l'autonomie. Ainsi, nous rejoignons Debaisieux (2009) selon laquelle l'intégration des corpus en classe constitue un défi pour la didactique du FLE.

Dans une perspective didactique, Fligelstone (1993) propose trois façons d'enseigner l'utilisation des corpus :

- (1) « teaching about » : cet usage consiste à « enseigner la linguistique de corpus et sa méthodologie » (Cavalla & Loiseau, 2014, p. 165), notamment pour des spécialistes des corpus tels que les talistes et les linguistes.
- (2) « teaching to exploit » : le deuxième usage est de former les apprenants à exploiter les corpus (par exemple, pour apprendre une langue étrangère dans le cas qui nous intéresse ici). L'objectif est d'amener l'apprenant à

« s'approprier l'outil d'exploitation du corpus afin de le manipuler à sa guise » (*ibid.*, p. 167).

- (3) « exploiting to teach » : cette approche vise à enseigner l'usage des corpus à des formateurs (enseignants, talistes ou linguistes) à des fins d'utilisation (pas forcément didactique). Concernant l'utilisation didactique, il s'agit d'« exploiter les corpus et l'interprétation des résultats pour enseigner une langue » (*ibid.*, p. 165).

De même que Cavalla (*ibid.*), nous adoptons les deux dernières approches dans les exploitations didactiques concrètes. Concernant « teaching to exploit », il est possible d'envisager deux manières de procéder : soit la manière indirecte où l'apprenant (et peut être l'enseignant aussi en tant que « l'utilisateur novice des corpus ») (Kübler, 2014) travaille sur des exercices conçus à partir de concordances présélectionnées ; soit la manière directe où l'apprenant est amené à découvrir comment interroger le corpus pour ensuite être capable de chercher seul des éléments linguistiques dans le corpus. Ici, notre réflexion porte notamment sur la manière indirecte de l'approche « teaching to exploit », parce que d'une part, le corpus Scientext « n'est pas a priori conçu pour une utilisation pédagogique » (Cavalla & Loiseau, 2014, p. 167), d'où l'importance de la sélection des données utilisables en classe ; d'autre part, comme expliqué dans le chapitre I, afin d'éviter à l'apprenant de se perdre dans des exemples trop fournis ou trop complexes et d'encourager un apprentissage progressif. Il est en effet important de passer par une phase d'initiation au cours de laquelle l'apprenant peut travailler sur des lignes de concordances sélectionnées et des exercices préparés par les enseignants (Chambers, 2010 ; Cavalla & Loiseau, 2014). Quant à l'approche « exploiting to teach », comme expliqué plus haut, nous montrons de quelle façon exploiter notre base de données de patrons (ressources des corpus) pour créer du matériel pédagogique. Ceci constitue un exemple illustrant nos réflexions sur l'enseignement des constructions verbales avec le corpus en FLE. Notons par ailleurs que ces deux approches accordent une place importante à la réflexion métalinguistique du fait que le travail sur les concordances met en jeu une activité d'observation et des compétences d'induction.



## 5.5.2. Deux expérimentations initiales

Ayant l'objectif d'enseigner les constructions verbales du LST avec le corpus (en particulier la consultation indirecte), nous avons proposé deux expérimentations didactiques auprès de deux publics différents d'apprenants. Ces premières expériences nous permettent de lier les analyses linguistiques aux pratiques concrètes et de mener une réflexion sur les apports et les limites des corpus dans l'enseignement des constructions verbales du LST.

### 5.5.2.1. Première expérimentation à l'Université de Xi'an

Au sein du département du français à l'Université des Études Internationales de Xi'an (Chine), la première expérimentation a été réalisée de mi-octobre à fin-novembre 2014 sur 6 séances (2h par séance) auprès de 16 étudiants chinois<sup>5</sup> de master 1 (niveau B1/B2) qui apprennent le français comme langue étrangère. L'objectif de cette expérimentation était d'amener les étudiants à se familiariser avec l'écrit scientifique et avec les constructions verbales liées à l'expression de l'opinion telles que *il est possible de considérer que, on peut admettre que, on conteste l'affirmation, etc.*

Les six séances ont été divisées en deux parties :

- (1) une séquence générale sur l'écrit scientifique qui amène l'apprenant à découvrir le genre de l'écrit scientifique
- (2) une séquence particulière sur les constructions verbales des verbes d'opinion pour entraîner l'apprenant à l'utilisation de constructions verbales spécifiques relevant de l'opinion (*considérer, affirmer, estimer, admettre, rejeter, etc.*)

Les exercices proposés dans cette expérimentation sont fournis dans l'Annexe 3. Notons par ailleurs que les exercices sont conçus à partir des données du corpus Scientext. Lors de la séquence générale sur l'écrit scientifique, nous avons proposé

---

<sup>5</sup> Ces étudiants ont participé volontairement à l'expérimentation. Pendant la première séance, il est apparu que les 16 étudiants avaient peu de connaissance sur l'écrit scientifique et une connaissance limitée sur les verbes d'opinion.

notamment des exercices de repérage (par exemple, repérer les différentes parties de l'article scientifique) pour que les étudiants puissent découvrir eux-mêmes la structure organisationnelle et le fonctionnement des phénomènes linguistiques de l'écrit scientifique. Quant à la séquence sur les verbes d'opinion, nous avons proposé une progression classique allant des exercices de repérage des constructions verbales des verbes d'opinion vers le réemploi des éléments étudiés dans une tâche rédactionnelle, en passant par l'approfondissement des savoirs relatifs aux propriétés sémantiques et syntaxiques, à travers des exercices d'association (sens et constructions correspondantes), des exercices lacunaires (verbes synonymes) et des exercices de reformulation.

Il faut souligner que les activités de repérage ont été privilégiées dans les séances parce que notre public était peu familiarisé avec l'écrit scientifique (il confond l'écrit journalistique et l'écrit scientifique) et moins encore avec les constructions verbales liées à l'expression de l'opinion. Pendant le repérage du phénomène, nous avons privilégié l'approche inductive qui se marie bien avec l'utilisation des corpus pour l'enseignement/apprentissage des phénomènes linguistiques. En effet, cette approche crée une certaine motivation chez les étudiants parce qu'ils découvrent eux-mêmes les régularités, le fonctionnement des phénomènes linguistiques. Par ailleurs, la formulation des règles par eux-mêmes développe chez eux une prise de conscience linguistique et cela faciliterait donc une bonne mémorisation. Cependant, nous n'avons pas rejeté l'approche déductive qui peut être utilisée pour apporter des explications et renforcer les connaissances.

À l'issue de l'expérimentation, un questionnaire a été proposé pour connaître les acquis et les avis des étudiants. De manière générale, les étudiants ont apprécié l'expérience qui leur a permis de 1) travailler sur des extraits d'article scientifique 2) connaître des constructions verbales variées liées à l'expression de l'opinion. Cependant, un tiers des étudiants a trouvé que les exemples sélectionnés étaient un peu difficiles puisque certains exemples évoquent des notions spécialisées inconnues. Deux étudiants ont regretté de ne pas avoir davantage d'activités de production écrite. Il faut souligner que les séances proposées constituaient notre première tentative d'exploitation indirecte des données du corpus et présentaient donc certaines limites. D'abord, le repérage du phénomène n'est pas évident et nous n'avons pas

suffisamment anticipé les difficultés des étudiants. Nous pensions qu'il serait profitable de privilégier les activités en petits groupes afin que les étudiants s'entraident. Ensuite, nous n'avons pas proposé suffisamment d'entraînement à l'écrit. Il est apparu que des activités rédactionnelles de courts paragraphes par imitation, dans lesquelles les étudiants pourraient réemployer les mots appris, pourraient constituer une étape importante pour qu'ils puissent ensuite se sentir plus à l'aise dans la rédaction de leurs écrits universitaires.

Passons maintenant à la deuxième expérimentation.

### 5.5.2.2. Deuxième expérimentation à l'Université Stendhal

La deuxième expérimentation s'est effectuée de mi-mars à mi-mai 2015 au sein de l'Université Stendhal (maintenant l'Université Grenoble Alpes) pendant une formation de Français sur Objectifs Universitaires (désormais FOU) dispensée auprès de six étudiants chinois de Master 2 en SHS<sup>6</sup>. L'objectif de cette expérimentation était de leur apporter une aide dans la rédaction de leurs mémoires de master, en les amenant à comprendre le fonctionnement de certaines structures langagières spécifiques à l'écrit scientifique. L'expérimentation s'est composée de six séances (3 heures par séance) dans lesquelles nous avons abordé les constructions verbales par champs sémantiques :

- (1) la construction de la problématique (*on s'intéresse à la question, nous formulons les questions suivantes*)
- (2) la prise de position par rapport à d'autres théories existantes (*on admet que, on rejoint le point de vue*)
- (3) le choix de l'approche et la collecte des données (*les données que nous avons recueillies, nous adoptons l'approche*)
- (4) l'interprétation à partir d'un constat ou d'une illustration (*on peut constater que, comme le montre le tableau*)
- (5) le rappel des points traités et la présentation des résultats (*l'analyse conduit à la conclusion, le résultat a confirmé l'hypothèse*)

---

<sup>6</sup> Les six étudiants chinois ont participé volontairement à l'expérimentation. Ils viennent de différentes spécialités dont la linguistique, les sciences de l'information et de la communication, les sciences du langage et le traitement automatique des langues.

- (6) l'utilisation du corpus au service de la rédaction des mémoires : rédiger la conclusion à l'aide des outils Dicorpus et Scribplus

L'organisation de ces séances suit ainsi la rédaction des mémoires, allant de la problématique à la présentation des résultats. Le découpage par champs sémantiques permet de regrouper les constructions verbales ayant des fonctions similaires, et par conséquent de « situer l'apprenant à un moment de son écriture » (Cavalla & Loiseau, 2014, p. 172). Chaque séance suit une progression classique et se compose des étapes suivantes :

- (1) Découverte du lexique et des constructions verbales : il est d'abord proposé des exercices de repérage sur un ou deux extraits qui introduisent les phénomènes linguistiques.
- (2) Analyse des phénomènes linguistiques : les étudiants travaillent en binôme sur plusieurs extraits pour repérer les phénomènes linguistiques, l'enseignant souligne par la suite dans le texte les éléments sur lesquels il veut qu'ils s'arrêtent spécifiquement afin de leur faire prendre conscience de ces éléments.
- (3) Fixation des connaissances : il est proposé de travailler sur les différents aspects liés aux verbes tels que la synonymie, les alternances syntaxiques, ou les valeurs sémantiques, afin de renforcer les acquis des étudiants.
- (4) Classement sémantique des constructions/routines : les étudiants doivent classer les constructions ou les routines en fonction de leurs valeurs sémantiques ou fonctions rhétoriques afin qu'ils comprennent à quoi servent les éléments linguistiques étudiés et qu'ils fassent un classement de ces éléments.
- (5) Réemploi : les étudiants sont amenés à produire un court paragraphe en employant les constructions verbales étudiées.

Afin de faciliter l'aide à l'écriture et de guider l'apprenant vers l'exploitation autonome des corpus, lors de la dernière séance, les étudiants ont été amenés à travailler sur des concordances présélectionnées dans l'outil Dicorpus et à rédiger un

court paragraphe relevant de la conclusion (le rappel des points traités et la présentation des résultats) à l'aide de l'outil Scribplus (Mangenot & Phoungsub, 2010). Ainsi que nous l'avons mentionné dans le Chapitre I, l'outil Dicorpus, conçu par Tran et Falaise (2016), fonctionne comme un dictionnaire mettant à la disposition des enseignants et des apprenants la description des caractéristiques syntaxiques et sémantiques du lexique LST ainsi que les exemples des concordances extraites. La figure 5-11 illustre l'interface de Dicorpus selon l'entrée onomasiologique.

Les verbes dans les écrits scientifiques   Conclusion			
MODE D'ACCÈS	POUR EXPRIMER...	EXPRESSION	
<ul style="list-style-type: none"> <li>▶ Accès par sens</li> <li>▶ Accès par expression</li> </ul>	... ... faire un rappel du travail présenté ... résumer l'essentiel des résultats	analyser (** fréquent moyen) étudier (** très fréquent) examiner (** fréquent moyen) explorer (** fréquent moyen) traiter (** fréquent moyen) atteindre (** fréquent moyen) élaborer (** très fréquent) établir (** très fréquent) réaliser (** très fréquent)	
<b>TYPE DE VERBES</b>			
verbes d'examen			
<b>ASSOCIATION AVEC ADVERBES</b>			
- ainsi, aussi   - ensuite   - notamment			
<b>CONSTRUCTIONS</b>			
nous avons analysé les résultats, ce travail permet d'analyser, il s'agit d'analyser, il est nécessaire d'analyser			
<b>SIGNIFICATION</b>			
Faire une analyse sur l'objet d'étude ou quelque chose d'abstrait (évolution, relation, processus, effet, résultat, modèle) en le décomposant en ses éléments de manière à le définir, le classer, le comprendre			
Exemples validés sur le corpus Scientext			
7	<p>Au final, la comparaison disciplinaire se présente donc non comme une fin en soi, mais comme une entrée pour l'analyse des textes dans leur diversité ; cette diversité nous semble primordiale pour échapper aux descriptions unifiantes et réductrices. Si l'intérêt d'<b>analyser</b> l'écrit de recherche au sein de genres est de montrer qu'on n'écrit pas la science de la même manière dans les différents champs, il faut aussi convenir qu'il en va de même au sein d'un seul champ. En d'autres termes et pour reprendre une expression dont nous avons montré l'importance dans les articles de LET le linguiste se doit de " résister à la tentation " d'une interprétation des textes transparente vis-à-vis des pratiques supposées, en l'occurrence les pratiques culturelles d'une discipline.</p> <p>[Linguistique - Thèse - Conclusion] L'article de Recherche en Sciences du Langage et en Lettres, Figures de l'auteur et identité disciplinaire du genre, Fanny Rinck</p>		
7	<p>Revenons cependant sur le problème du statut du scripteur. A travers cette variable, nous avons cherché à <b>analyser</b> un degré d'acculturation au genre. Deux types d'interprétations de cette idée d'acculturation peuvent être envisagés au vu de la spécificité des textes de doctorants par rapport à ceux des non-doctorants. D'abord, l'idée d'un degré d'acculturation se base sur celle d'un degré de connaissance du champ et de ses enjeux ; c'est l'hypothèse interprétative avancée face aux modes contrastés de renvoi aux sources et à la faible présence de noms de courant dans les articles de doctorants. La connaissance du champ ne suffit cependant pas à elle seule à expliquer cet usage moindre : le scripteur doit se sentir légitimé à utiliser de telles étiquettes, qui en retour contribuent à une posture légitimante de surénonciation. Cette hypothèse revient à avancer que les pratiques des néophytes sont amenées à évoluer dans le sens des pratiques des non-doctorants.</p> <p>[Linguistique - Thèse - Conclusion] L'article de Recherche en Sciences du Langage et en Lettres, Figures de l'auteur et identité disciplinaire du genre, Fanny Rinck</p>		
7	<p>Nous avons <b>analysé</b> certains de nos résultats à la lumière de la théorie " <i>Frame, then Content</i> " (MacNeilage, 1998) qui définit le cadre syllabique comme un simple cycle de l'oscillation mandibulaire, le début et la fin de chaque syllabe étant déterminés par un maximum de fermeture (position haute de la mandibule). Le cycle mandibulaire fournit la contrainte temporelle de la syllabe (période de 1/4 de seconde environ) qui se traduit par une constante de durée (Greenberg, 1998), et la contrainte mécanique de mode, les articulations vocaliques et consonantiques s'inscrivant respectivement dans les phases d'ouverture et de fermeture de la cavité buccale. La théorie " <i>Frame, then Content</i> " répond du même coup à deux grands principes de la phonologie de la syllabe (Redford, 1999 ; Redford et al., 2000) :</p> <p>[Linguistique - Thèse - Conclusion] Structures syllabiques et lexicales des langues du monde. Données, typologies, tendances universelles et contraintes substantielles, Isabelle Rousset</p>		

**Figure 5-11 : Entrée onomasiologique de l'outil Dicorpus\_ 'faire un rappel du travail présenté'**

En plus de l'outil Dicorpus, le logiciel Scribplus fournit une aide supplémentaire à l'écriture. Il accompagne les étudiants dans leur rédaction en proposant une série de questions. En répondant à chaque question, les étudiants peuvent écrire des phrases à partir desquelles sera construit un énoncé final (Phoungsub, 2012) (voir la figure 5-12).

Genre/Type: argumentatif

Thème: Rédiger la conclusion pour votre mémoire

Étape2: Vous allez tout d'abord revoir la partie d'introduction rédigée (ou le plan de votre mémoire), et puis faire un rappel de l'étude aux lecteurs en reformulant le sujet de recherche, la problématique ou la méthodologie adoptée. Vous pouvez aussi vérifier si l'étude est menée comme prévu ou si l'étude est bien réalisée. Développez-les en quatre lignes et n'oubliez pas d'utiliser les constructions étudiées.

Figure 5-12 : L'activité rédactionnelle sur l'outil Scribplus

Quatre étapes/questions s'articulent dans cette activité rédactionnelle, incitant les étudiants à employer les constructions verbales étudiées. Les fiches pédagogiques de

---

cette expérimentation sont présentées dans l'Annexe 4.

De l'analyse des questionnaires de satisfaction distribués à la fin de l'expérimentation, ressort un bilan plutôt satisfaisant. Tous les étudiants ont trouvé les constructions verbales abordées bien utiles pour la rédaction de leur mémoire, étant « agréablement surpris de ce qu'ils ont appris à travers une telle expérimentation »<sup>7</sup> et ils comptent réinvestir les éléments appris dans leurs écrits universitaires. Deux étudiants sont plus réticents à l'utilisation de l'outil Dicorpus. D'une part, il apparaît que nous n'avons pas bien expliqué l'intérêt du corpus en raison d'un temps limité ; d'autre part, ces deux étudiants ont un style d'apprentissage plutôt déductif et préfèrent avoir des listes d'expressions à retenir. Il faudrait alors vérifier si avec un enseignement en listes d'expressions à apprendre par cœur, ils réussiraient tout autant à rédiger. En tant que didacticienne, nous pensons que l'apprentissage en liste ne donne pas de clés pour l'utilisation en contexte.

Ces constats rejoignent ceux mentionnés par Lewis (2006) et Boulton (2008) sur le lien étroit entre une approche sur corpus et les styles d'apprentissage des apprenants : « Les résultats indiquent clairement que les styles d'apprentissages [inductif ou déductif]... ont une incidence sur l'efficacité d'une approche sur corpus... [qui] favorise fortement ceux d'un style inductif » (Lewis, 2006, p. 104). Dans le travail de Lewis (2006), 78% des apprenants sont considérés comme déductifs et préfèrent avoir à leur disposition des règles pour un travail sur corpus. L'enseignement doit non seulement tenir compte du niveau et du besoin des apprenants, mais également de leur style d'apprentissage. L'approche déductive s'avère nécessaire pour réduire la surcharge cognitive des étudiants et pour ne pas décourager ceux qui ont un style inductif. Comme nous l'avons évoqué ci-dessus, l'approche inductive nous semble un type d'entrée qui correspond bien à l'utilisation du corpus selon le DDL, pourtant les apprenants peuvent choisir leur propre approche.

De manière générale, nous pensons que les étudiants ont tiré profit de ces activités pédagogiques, car pendant les cours ils ont développé progressivement leurs connaissances sur les constructions verbales du LST en faisant preuve de leur capacité à repérer, à classer et à conceptualiser les règles de fonctionnement de l'écrit

---

<sup>7</sup> Réponse d'une des étudiants qui a participé à notre expérimentation.

scientifique. En examinant les écrits des étudiants, nous avons constaté qu'ils avaient su réemployer la plupart des constructions étudiées dans un contexte approprié.

Pour conclure, ces deux expérimentations nous ont permis de prendre du recul et de mener une réflexion sur l'enseignement des éléments phraséologiques au moyen des corpus. Nous résumons ci-dessous les idées qui pourraient guider des expérimentations didactiques ultérieures ou d'autres pratiques d'enseignement.

Sur le plan des démarches adoptées, il nous semble important d'associer le lexique scientifique transdisciplinaire et l'approche déductive, en prenant en compte les différents styles d'apprentissages des étudiants. L'approche inductive semble bien se marier avec l'utilisation des corpus mais elle présente certaines limites, notamment lorsque l'apprenant n'est pas bien guidé ou lorsqu'il est habitué au raisonnement déductif. Par ailleurs, il faut relier davantage la phraséologie et la recherche sur corpus, mettant en avant l'utilité des corpus. Dans les deux expérimentations, nous n'avons proposé que la consultation indirecte du corpus et la question de la recherche sur corpus n'a pas été traitée. Cependant, nous pensons que l'enseignement des phénomènes phraséologiques avec les corpus doit se fixer comme objectif final d'apprendre à apprendre, c'est-à-dire qu'il faut guider l'apprenant à maîtriser l'utilisation des divers corpus pour qu'il puisse chercher et travailler à sa manière sur les éléments qui lui sont directement utiles. Pour ce faire, nous rejoignons le point de vue de Boulton (2015) selon lequel il est profitable de partir de ce que connaissent déjà les apprenants (ce qui est le fondement du socio-constructivisme de Vygostky). La première étape de travail pourra commencer par une recherche sur les outils que maîtrisent les étudiants, tels que Google Scholar, google advance search, Ngram ou Linguee, avant de leur demander de faire la recherche sur des corpus spécialisés comme Scientext. En faisant la comparaison sur différents corpus, les étudiants pourront apprécier l'intérêt du corpus Scientext pour leurs écrits scientifiques.

Sur le plan des exercices proposés, les résultats des expérimentations nous ont confirmé que le repérage du phénomène est essentiel pour sensibiliser les apprenants aux phénomènes phraséologiques. En revanche, il nous semble important de varier les modalités des exercices pour éviter la répétition et maintenir leur attention. En effet, nous avons remarqué que les exercices de repérage proposés sont trop répétitifs et que



les étudiants s'ennuient facilement au bout de trois extraits. Par ailleurs, au vu des difficultés en matière de terminologie linguistique dans le corpus Scientext, nous pensons qu'avant de travailler sur les textes du corpus Scientext, il est nécessaire de passer par une étape préliminaire, dans laquelle l'enseignant peut choisir et donner des exemples venant de l'écrit scientifique de vulgarisation orientés vers le grand public (par exemple des revues encyclopédiques, et pourquoi pas de construire un corpus de vulgarisation scientifique comme Sciences Humaines<sup>8</sup>).

À travers ces deux expérimentations, nous avons montré que l'enseignement des constructions verbales du LST au moyen du corpus Scientext présente un grand intérêt. La ressource lexicale que nous avons proposée fournit un matériau pédagogique prêt à l'emploi, permettant de faire prendre conscience à l'apprenant les spécificités des verbes dans l'écrit scientifique. Ici, le corpus Scientext joue un rôle important, qu'il s'agisse soit d'une mise en place des exercices variés soit d'une exploitation des textes authentiques (Cavalla & Loiseau, 2014). Les exercices ainsi que les outils Dicorpus et Scribplus que nous avons mis en place contribuent ensemble à faciliter la rédaction universitaire pour ces étudiants chinois, en répondant à leurs besoins linguistiques. Notre proposition didactique pourrait être appliquée à d'autres contextes d'enseignement mais il revient à l'enseignant de choisir et d'adapter la démarche et les outils à son public en fonction des besoins spécifiques de ce dernier.

En suivant ces réflexions didactiques par rapport aux deux expérimentations réalisées, nous proposons deux exemples d'activités dans lesquelles nous tâchons d'éviter les écueils relevés précédemment. Le premier exercice consiste à repérer le verbe *montrer* et les constructions verbales qui y sont associées. L'exercice s'effectue en groupe. Le travail en groupe permet aux apprenants de construire ensemble un savoir et d'éviter les échecs si l'un d'entre eux n'arrive pas bien à suivre. Par ailleurs, les apprenants ne se retrouvent pas seuls face à des éléments nouveaux, ce qui favorise ainsi l'argumentation entre eux et la réflexion sur les éléments nouveaux. Le deuxième exercice vise à amener les apprenants à faire des listes récapitulatives à la fin de l'étude des constructions verbales afin d'aider à consolider leur mémorisation.

---

<sup>8</sup> Ce corpus est disponible en ligne : <https://www.ortolang.fr/market/corpora/scienceshumaines>, consulté le 12 décembre 2016.

### 5.5.3. Exemples d'exercices adaptés

Les deux exemples d'activités que nous souhaitons mettre en avant présentent deux types d'exercices qui ont bien fonctionné lors de nos expérimentations. Nous les avons élaborés en prenant en compte nos réflexions et en faisant un lien avec les résultats d'analyses linguistiques du Chapitre IV. Rappelons que les exercices proposés ont pour objectif principal d'amener les étudiants progressivement vers une palette plus large des constructions verbales caractéristiques des pratiques expertes, et que celles-ci sont construites en fonction des besoins linguistiques des étudiants chinois et dans la perspective d'associer l'enseignement des phénomènes phraséologiques à l'utilisation des corpus en classe de FLE.

#### 5.5.3.1. Repérage réflexif

Comme expliqué plus haut, l'activité de repérage dans les textes est essentielle pour sensibiliser l'apprenant aux phénomènes phraséologiques. L'enseignant peut utiliser les corpus pour extraire des concordanciers pertinents autour d'un phénomène phraséologique. Dans le tableau 5.7 figure un exemple du concordancier du verbe *montrer*, extrait du corpus Scientext. Ici, nous proposons de faire observer aux étudiants les structures syntaxiques (variations syntaxiques, temps et mode du verbe, etc.) pour relever par la suite les propriétés sémantiques du verbe.

Consigne : En binôme, observez les phrases et répondez aux questions suivantes.

Contexte gauche	Occurrences	Contexte droit
Les exemples	montrent	que ces motivations répondent à des rapports approximatifs ou des similarités partielles.
L'analyse des discours	montre	que l'expertise la plus développée est conduite par les ayants droit qui, pour caractériser le milieu et expliquer son usage, prennent en compte des paramètres plus complexes que ceux des nomades originaires de régions lointaines.
Les données	montrent	que la rencontre virtuelle passe plus souvent par des forums ou des groupes de discussion (pour 32 individus sur 82) ou par des sites de "chat" ou de messagerie instantanée (pour 34 individus) que par des sites de rencontre (13 individus) ou des sites de jeux (13 individus).
Par ailleurs si, comme on l'a	montré	, sa participation au conseil municipal est tout aussi disqualifiante que valorisante, il faut rappeler que le titre d' élu peut fonctionner comme une ressource dans des sphères éloignées de la scène communale.
Premièrement, comme	montré	, la définition même de cette nature n'est rendue possible que par

nous l'avons		une analyse de son rapport avec l'Homme à différentes échelles de temps et d'espace.
Dans la situation de libre entrée envisagée, nous	montrons	que si le coût marginal de contrôle est suffisamment faible, le prix de la haute qualité, à l'équilibre avec test, est supérieur au coût marginal, mais il est inférieur au prix de la haute qualité à l'équilibre sans test.
On peut	montrer	que l'on peut trouver de manière équivalente la valeur de ce revenu seuil en comparant la fonction d'utilité indirecte dérivée de ce modèle à la fonction d'utilité indirecte issue du modèle contraint où le niveau de la taxe serait nul.

**Tableau 5.7 : Concordancier du verbe *montrer***

L'enseignant pourrait poser les questions suivantes afin de guider l'apprenant et de favoriser l'interprétation :

- Syntaxe : Quel est l'élément le plus fréquent avant et après le verbe *montrer* ? Si vous connaissez des formes figées avec le verbe *montrer*, est-ce que vous en repérez dans ces phrases ? L'enseignant peut guider l'apprenant pour le repérage de la construction en *comme* et la construction en complétive *que*.
- Lexique : quel est le sujet du verbe *montrer* ? L'enseignant peut mettre l'accent sur le sujet du verbe *montrer* qui est soit humain soit non-humain. Quand le sujet est humain, il s'agit souvent des pronoms *on* ou *nous*.
- Sens : est-ce que *montrer* a ici plusieurs sens ? Est-ce que les formes figées que vous avez repérées ont des sens particuliers ? L'enseignant peut demander à l'apprenant de s'arrêter sur la construction en *comme* et guider l'apprenant à cerner le sens ou la fonction.

### 5.5.3.2. Classification sémantique

Un autre type d'exercice que nous souhaitons mettre en avant consiste à classer les constructions verbales selon leurs valeurs sémantiques ou fonctions rhétoriques. Nous supposons que l'apprenant les a déjà travaillées en amont. Il permet à l'apprenant de mieux comprendre à quoi servent les constructions étudiées et de produire sa propre liste lexicale. Dans le tableau 5.8, nous proposons de classer les formules suivantes dans trois cases différentes selon leurs fonctions rhétoriques. Il s'agit de trois groupes sémantiques appartenant au champ sémantique du

positionnement.

Consigne : En binôme, classez les formules suivantes dans le tableau ci-après. Ajoutez-y les synonymes récoltés.

1. Faire l'hypothèse que
2. Notre première prédiction est que
3. On peut admettre que
4. Il en résulte que
5. Ce constat amène à considérer que
6. L'hypothèse prend son essor en
7. Il est possible d'affirmer que
8. On peut en déduire que
9. On peut postuler que
10. Ces résultats peuvent être interprétés comme

Exprimer une opinion	Présenter une hypothèse	Proposer une interprétation

**Tableau 5.8 : Exercice de classement sémantique**

Exprimer une opinion	Présenter une hypothèse	Proposer une interprétation
On peut admettre que	Faire l'hypothèse que	Il en résulte que
Il est possible d'affirmer que	L'hypothèse prend son essor en	On peut en déduire que
Ce constat amène à considérer que	Notre 1re prédiction est que	Ces résultats peuvent être interprétés comme
	Notre seconde prédiction est que	
	On peut postuler que	

**Tableau 5.9 : Correction de l'exercice**

Pour résumer, sur le plan didactique, afin de remédier aux lacunes langagières en matière de constructions verbales du LST, nous avons présenté deux types de ressources destinées au public du FLE. Ces ressources démontrent bien l'apport des corpus à l'enseignement des phénomènes phraséologiques et représentent nos premières réflexions didactiques dans la perspective d'intégrer les corpus dans les pratiques enseignantes. Dans une recherche ultérieure, nous envisageons d'enrichir notre ressource lexicale des patrons, de mettre à l'épreuve et de varier nos exercices dans d'autres contextes d'enseignement.

## CONCLUSION GENERALE

Cette recherche met en jeu à la fois le TAL, la linguistique de corpus, la didactique et la lexicographie qui sont reliés par le même objet d'étude – les phénomènes phraséologiques, plus précisément, les patrons lexico-syntaxiques des verbes du LST. Le propos principal de ce travail de recherche a été de mettre en évidence les propriétés sémantiques et syntaxiques des constructions verbales scientifiques. En mettant en relief les rôles que jouent ces constructions dans l'écrit scientifique et en ciblant les besoins langagiers des apprenants chinois, nous avons souligné leur importance dans l'enseignement et l'apprentissage de l'écrit académique. En guise de conclusion, nous allons retracer les résultats essentiels de notre travail avant de prendre du recul et de proposer des perspectives pour de futures études.

### Résultats essentiels

D'abord, du point de vue méthodologique, la combinaison de méthodes du TAL et de l'analyse manuelle a permis de modéliser les constructions verbales récurrentes en intégrant dans les patrons des informations syntaxiques, sémantiques, et lexicales. Il a été démontré qu'en s'appuyant sur les propriétés sémantiques et syntaxiques des verbes, il était possible de saisir les différents sens du verbe et de proposer une classification sémantique.

Ensuite, l'étude des patrons verbaux a mis en évidence des spécificités liées au comportement lexico-syntaxique des verbes dans l'écrit scientifique, qui se manifestent notamment à travers les constructions syntaxiques récurrentes, les sens mobilisés et les fonctions associées. En examinant les verbes d'analyse et les verbes de positionnement, nous avons fait émerger des propriétés propres à chaque groupe de classe. Les verbes d'analyse, qui relient le rôle du chercheur aux objets d'étude, remplissent ainsi des valeurs métatextuelles importantes ; tandis que les verbes de positionnement, employés pour exprimer des prises de position marquées ou discrètes de l'auteur, sont davantage associés aux valeurs argumentatives. Les résultats d'analyse obtenus se situent dans la continuité d'autres travaux effectués au sein du

Lidilem (Boch & Rinck, 2010 ; Rinck, 2010a ; Tutin & Grossmann, 2014), selon lesquels l'écrit scientifique est une écriture à visée objectivante et argumentative. D'une part, la construction des faits scientifiques s'appuie sur les exemples, les données et les résultats ; d'autre part, l'élaboration des connaissances nouvelles se fonde sur l'expression des opinions et la justification des choix.

Sur le plan didactique, la ressource lexicale des patrons verbaux présentée comme un outil d'aide à l'écrit académique représente un moyen d'initier les étudiants étrangers aux caractéristiques du genre. Le traitement à la fois sémasiologique et onomasiologique a permis d'associer les formes lexicales et les aspects sémantiques aux fonctions rhétoriques, celles-ci sont « jugées complémentaires dans une perspective didactique » (Rinck, 2010a, p. 439).

Il convient de souligner un autre apport principal de cette recherche qui relève du diagnostic des besoins langagiers des apprenants chinois. En établissant une grille de diagnostic centrée à la fois sur les erreurs et les cas de suremploi et de sous-emploi de verbes du LST et de constructions verbales, nous avons montré que les difficultés des apprenants chinois sont principalement liées aux lacunes discursives (et non seulement lexicales). En effet, le statut de novice en matière de rédaction scientifique confronte, de manière identique les étudiants natifs et non-natifs, aux difficultés liées au genre. Ces difficultés les empêchent d'assumer le rôle de chercheur/argumentateur, d'étayer leur position par rapport aux travaux d'autrui, de mettre en avant leurs objets d'étude, de construire la cohérence textuelle et de dérouler leur raisonnement. Les erreurs lexicales et grammaticales pourraient diminuer au fur et à mesure de l'apprentissage de la langue, mais les difficultés au niveau discursif nécessitent un accompagnement pédagogique spécifique. Ces difficultés présentent ainsi un enjeu didactique.

Pour répondre à ces besoins langagiers, nous avons proposé des activités d'enseignement des verbes du LST en intégrant un corpus spécialisé d'écrits scientifiques en classe de langue. Les activités mises en place ont montré comment faire travailler les apprenants sur les différents aspects des verbes (associations lexicales, structures syntaxiques). Un tel enseignement présente le grand intérêt de sensibiliser les apprenants aux formules discursives, aux processus d'analyse ou au

positionnement scientifique, permettant de former une représentation sur le discours scientifique.

Enfin, cette recherche inscrite dans la linguistique de corpus a confirmé l'intérêt des corpus à plusieurs égards :

1. pour la construction des ressources lexicales
2. pour le diagnostic : repérage des difficultés chez les apprenants
3. pour l'enseignement, par la médiation de l'enseignant : repérage d'exemples, utilisation de concordances sélectionnées

### Limites et perspectives

Bien que cette étude ait permis de mieux comprendre les régularités sémantiques et syntaxiques des verbes du LST et les besoins langagiers des apprenants, elle connaît plusieurs limites. La première limite réside dans la phase d'extraction automatique des cadres. Il nous semble important de gérer les cas d'erreur d'analyse syntaxique (verbe sans sujet, transitif sans objet) et d'améliorer les regroupements d'occurrences en vue de proposer un calcul approximatif de fréquence pour chaque patron. Nous envisageons également d'intégrer des informations sémantiques automatiquement dans les cadres de sous-catégorisation. Par exemple, le verbe *démontrer* sélectionne des sujets de la classe [[qualité|relation]] : *importance, effet, existence, capacité, limite, rôle*.

La deuxième limite concerne le petit nombre de verbes traités. Cette thèse a permis de dégager des patrons lexico-syntaxiques et une trentaine de fonctions rhétoriques associées à 32 verbes. La perspective ouverte de cette recherche consiste d'abord à intégrer dans cette ressource des patrons les verbes du LST relevant d'autres classes sémantiques : les verbes de communication, les verbes de relation, etc. Afin de ne pas être limité aux verbes du LST, une autre piste intéressante est d'enrichir la ressource avec d'autres marqueurs linguistiques utilisés dans différentes fonctions rhétoriques, par exemple, les verbes polylexicaux comme *mettre en évidence*, ou les collocations comme *faire des hypothèses*. Notre objectif à long terme consiste à mettre cette ressource en ligne en l'associant à l'outil Dicorpus.

La troisième limite est liée à la taille et à la représentativité de notre corpus d'apprenants. Afin de tester si les résultats de diagnostic sont également valables pour d'autres publics, il est indispensable de constituer un corpus d'apprenants plus important avec des textes non corrigés. Nous comptons collecter prochainement des textes académiques auprès des étudiants sinophones du centre universitaire d'études françaises (CUEF) de l'Université Grenoble Alpes. Ces données nous permettront de réajuster notre grille de diagnostic.

Enfin, les deux expérimentations effectuées nous permettent de relever certains problèmes liés au principe même de l'enseignement avec corpus numérique. Concernant l'utilisation du corpus, les questions suivantes se posent : quel extrait de texte choisir pour l'enseignement ? Quelle place donner au corpus ? En fait, les exemples tirés du corpus spécialisé Scientext s'avèrent difficiles pour notre public d'étudiants (niveaux B1 et B2). Il est possible d'envisager une expérimentation dans laquelle on pourrait donner des exemples venant d'un corpus de vulgarisation scientifique en Sciences Humaines. La mise en place d'un corpus de vulgarisation scientifique permettrait de sensibiliser les apprenants aux régularités langagières sans les confronter à la question de la terminologie. Par la suite, on peut introduire un corpus spécialisé comme Scientext.

Si dans l'enseignement avec corpus, la prise de conscience des phénomènes linguistiques constitue la première étape, l'appropriation de l'outil corpus est une fin méthodologique. Il s'agit d'un apprentissage guidé par l'enseignant et qui vise à l'autonomie. En complément des outils classiques comme les dictionnaires, le corpus constitue une aide supplémentaire à la rédaction académique. Nous pensons avoir contribué à montrer l'utilité de l'outil corpus dans l'enseignement et l'apprentissage de l'écrit académique.



---

## BIBLIOGRAPHIE

- Ädel, A. (2003). *The Use of Metadiscourse in Argumentative Writing by Advanced Learners and Native Speakers of English*. (Ph.D. dissertation). Göteborg, Sweden : Göteborg University.
- Ädel, A. (2006). *The use of metadiscourse in argumentative texts by advanced learners and native speakers of English*. Amsterdam & Philadelphia John Benjamins Publishing Company.
- Aït-Mokhtar, S., Chanod, J.-P., & Roux, C. (2002). Robustness beyond shallowness : incremental dependency parsing. *Natural Language Engineering*, 8(2/3), 121-144.
- Anctil, D. (2005). *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux* (Mémoire de Master 1). Montréal : Université de Montréal.
- Anctil, D. (2010). *L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3<sup>e</sup> secondaire et description du rapport à l'erreur lexicale d'enseignants de français* (Thèse de doctorat). Montréal : Université de Montréal.
- Aston, G. (1997). Small and large corpora in language learning. *PALC 97 : Paractical applications in language corpora*, 51- 62.
- Azzopardi, S. (2010). Présentation : La linguistique « de » corpus au-delà des champs disciplinaires : questions et enjeux transversaux. *Cahiers de praxématique*, (54- 55), 11- 24.
- Bak, M., & Novakova, I. (2013). Le raisonnement dans les textes scientifiques : le cas des verbes causatifs. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours*. *Autour de Scientext*. Rennes : Presses Universitaires de Rennes, 101- 122.
- Bally, C. (1909 = 1951). *Traité de Stylistique française*, vol. 1. Genève : Librairie Georg and Cie ; Paris : Librairie Klincksieck, 3<sup>e</sup> édition.
- Béjoint, H. (2007). Informatique et lexicographie de corpus : les nouveaux dictionnaires. *Revue française de linguistique appliquée*, XII(1), 7- 23.
- Berthelot, J.-M. (Éd.). (2003). *Figures du texte scientifique*. Paris : Presses

---

universitaires de France.

- Besse, H., & Porquier, R. (1991). *Grammaires et didactiques des langues*. Paris : Didier.
- Biber, D. (1993). Using register-diversified corpora for general language studies, *19*(2), 219-241.
- Binon, J., & Verlinde, S. (2002). Comment le FOS s'est profilé dans la didactique du FLE : de la formation d'enseignants à l'université à la conception d'un dictionnaire du français des affaires. *Cahiers de l'ASDIFLE*, (14), 17-37.
- Bizzell, P. (1992). *Academic discourse and critical consciousness*. Pittsburgh : University of Pittsburgh Press.
- Boch, F. (2013). Former les doctorants à l'écriture de la thèse en exploitant les études descriptives de l'écrit scientifique. In F. Komesu et L. Tenani (Éd.) *Ecriture et discours, Revista Linguagem em (Dis)curso*, 3(13), 543-568.
- Boch, F., & Grossmann, F. (2002). Se référer au discours d'autrui, quelques éléments de comparaison entre experts et néophytes. *Enjeux, L'écrit dans l'enseignement supérieur*, II(54), 41-51.
- Boch, F., & Rinck, F. (2016). Anaphores démonstratives dans les écrits d'étudiants de Master. Comparaison avec les pratiques expertes. *Linx*, (72), 111-130.
- Boch, F., & Rinck, F. (2010). Pour une approche énonciative de l'écrit scientifique. *Lidil, Énonciation et rhétorique dans l'écrit scientifique*, (41), 5-14.
- Bolly, C. (2007). La notion de compétence « phraséologique » : Verbes à haute fréquence et corpus d'apprenants anglophones avancés. In I. González Rey (Éd.), *Les expressions figées en didactique des langues étrangères*. Cortil-Wodon : E.M.E. & InterCommunications, 33-51.
- Bolly, C. (2008). *Les unités phraséologiques : un phénomène linguistique complexe ? Séquences (semi-) figées avec les verbes prendre et donner en français écrit L1 et L2. Approche descriptive et acquisitionnelle* (Thèse de doctorat). Belgium : Université Catholique de Louvain.
- Bolly, C. (2011). *Phraséologie et collocations. Approche sur corpus en français L1 et L2*. Bruxelles : P.I.E./ Peter Lang.
- Boulton, A. (2015). Applying data-driven learning to the web. In A. Leńko-Szymańska & A. Boulton (Éd.), *Multiple Affordances of Language Corpora for Data-driven Learning*. Amsterdam : John Benjamins, 267-295.
- Boulton, A., Canut, E., Guerin, E., Parisse, C., & Tyne, H. (2013). Corpus et appropriation de L1 et L2. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (68-69), 9-32.
- Cadet, L., & Rinck, F. (2014). Des écrits de la formation à la didactique de l'écriture. *Le français aujourd'hui*, 184(1), 3-13.

- 
- Callies, M., & Zaytseva, E. (2013). The Corpus of Academic Learner English (CALE) – A new resource for the study and assessment of advanced language proficiency. In S. Granger, G. Gilquin, & F. Meunier (Éd.), *Twenty Years of Learner Corpus Research: Looking back, Moving ahead*. Louvain-la-Neuve : Presses universitaires de Louvain.
- Catach, N., Duprez, D., & Legris, M. (1980). *L'enseignement de l'orthographe : l'alphabet phonétique international, la typologie des fautes, la typologie des exercices*. Paris : Nathan.
- Cavalla, C. (2007). Réflexion pour l'aide à l'écrit universitaire auprès des étudiants étrangers entrant en Master et Doctorat. In J. Goes & J.-M. Mangiante (Éd.), *L'accueil des étudiants étrangers dans les universités francophones : sélection, formation et évaluation*. Arras : Artois Presses Université, 37-48.
- Cavalla, C. (2008). Les collocations dans les écrits universitaires : un français spécifique pour les apprenants étrangers. In O. Bertrand & I. Schaffner (Éd.), *Apprendre une langue de spécialité : enjeux culturels et linguistiques*. Paris : Editions Ecole Polytechnique, 93-104.
- Cavalla, C. (2009). La phraséologie en classe de FLE. *Les Langues Modernes*, (1). Consulté le 25 mai 2016 : <http://www.aplv-languesmodernes.org/spip.php?article2292>
- Cavalla, C. (2010a). Les écrits universitaires des étudiants étrangers : quelles normes présenter ? In O. Bertrand & I. Schaffner (Éd.), *Quel français enseigner ? La question de la norme dans l'enseignement/apprentissage*. Paris : Editions Ecole Polytechnique, 203-214.
- Cavalla, C. (2010b). Méthodologie d'apprentissage de l'écrit universitaire. *Le Français dans le Monde. Recherches et Applications, Faire des études supérieures en langue française*, (47), 153-161.
- Cavalla, C. (2014). Collocations transdisciplinaires d'apprentissage : réflexion pour l'enseignement. Dans I. Gonzalez-Rey (Éd.), *Outils et méthodes d'apprentissage en phraseodidactique*. Editions Modulaires Européennes, 151-169.
- Cavalla, C. (2015). Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE. *Linx*, (72), 95- 110.
- Cavalla, C., & Loiseau, M. (2014). Scientext comme corpus pour l'enseignement. In F. Grossmann & A. Tutin (Éd.), *L'Écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presse universitaire de Rennes, 163-180.
- Chambers, A. (2010). L'apprentissage de l'écriture en langue seconde à l'aide d'un corpus spécialisé. *Revue française de linguistique appliquée*, XV(2), 9-20.
- 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 2). Grenoble : Université Grenoble Alpes.
- Chen, P.-C., Hu, K.-C., & Ho, J.-Y. (2009). A study of academic vocabulary used in the abstracts of business and management journals. *Taiwan International ESP*

---

*Journal*, 1(1), 51-76.

- Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14(2), 30-49.
- Chollet, I., & Robert, J.-M. (2007). *Les verbes et leurs prépositions*. CLE international.
- Chuang, F.-Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora*, 1(2), 251-271.
- Ciekanski, M. (2014). Les corpus : de nouvelles perspectives pour l'apprentissage des langues en autonomie ? *Recherches en Didactique des Langues et Cultures : les Cahiers de l'acedle*, 11(1), 111-135.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Éd.), *Cambridge Handbook of English Corpus Linguistics* Cambridge : Cambridge University Press, 478-497.
- Cori, M., & David, S. (2008). Les corpus fondent-ils une nouvelle linguistique ? *Langages*, 171(3), 111-129.
- Coxhead, A. (2002). The Academic Word List : A Corpus-based Word List for Academic Purposes. In *Teaching and Language Corpora (TALC) 2000 Conference proceedings*, 73-89.
- Coxhead, A. (2010). A new Academic Word List. *TESOL Quarterly*, 34(2), 213-238.
- De Cock, S. (2014). Investigating typical features of spoken interactions in the Louvain International Database of Spoken English Interlanguage (LINDSEI). Attitudinal stance in the spotlight. Présenté à *Corpus Linguistics Group*, University of Oslo (03/02/2014).
- De Cock, S., & Tyne, H. (2014). Corpus d'apprenants et acquisition des langues. *Recherches en Didactique des Langues et Cultures : les Cahiers de l'acedle, L'association des chercheurs et enseignants en didactique des langues étrangères*, 11(1), 137-168.
- Debaisieux, J.-M. (2009). Des documents authentiques oraux aux corpus : un défi pour la didactique du FLE. *Mélanges CRAPEL*, (31), 35-56.
- Deschepper, C., & Thyron, F. (2008). L'entrée dans le supérieur et l'accès aux discours universitaires : opérationnaliser la notion de rapport à l'écrit dans un projet de formation. In S.-G. Chartrand & C. Blaser (Éd.), *Le rapport à l'écrit : Un outil pour enseigner de l'école à l'université*. Namur : Presses universitaires de Namur, 61-86.
- Di Vito, S. (2013). L'utilisation des corpus dans l'analyse linguistique et dans l'apprentissage du FLE. *Linx*, (68-69), 159-176.
- Drouin, P. (2007). Identification automatique du lexique scientifique transdisciplinaire. *Revue française de linguistique appliquée*, Vol. XII(2), 45-64.
- Dubois, J., & Dubois-Charlier, F. (1997). *Les Verbes français*. Larousse.

- 
- EL Maarouf, I. (2011). *Formalisation de connaissances à partir de corpus : modélisation linguistique du contexte pour l'extraction automatique de relations sémantiques* (Thèse de doctorat). Université de Bretagne Sud. Consulté le 16 janvier 2016 : <https://tel.archives-ouvertes.fr/tel-00657708/document>
- Falaise, A. (2014). Exploitation linguistique de corpus arborés d'écrits scientifiques à l'aide du logiciel ScienQuest. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presses Universitaires de Rennes, 123-143.
- Falaise, A., Tutin, A., & Kraif, O. (2011a). Exploitation d'un corpus arboré pour non spécialistes par des requêtes guidées et des requêtes sémantiques. In *Actes de la 18<sup>e</sup> conférence annuelle sur le Traitement Automatique des Langues Naturelles*. Montpellier, France. Consulté le 12 décembre 2016 à l'adresse <http://pro.aiakide.net/publis/2011TALNPaper-Falaise-Tutin-Kraif.pdf>
- Falaise, A., Tutin, A., & Kraif, O. (2011b). Une interface pour l'exploitation de corpus arborés par des non informaticiens : la plate-forme ScienQuest du projet Scientext. *TAL*, 52(3), 241-246.
- Fifielska, E. (2015). *Les constructions syntaxiques de l'écrit scientifique : exploration et analyses de corpus* (Mémoire en Sciences du langage). Grenoble : Université Grenoble Alpes. Consulté le 18 juin 2015 : <http://dumas.ccsd.cnrs.fr/dumas-01213405>
- Fillmore, C. J., Paul, K., & Mary, C. O. (1988). Regularity and idiomaticity in grammatical constructions : The case of let alone. *Language* 64(3), 501-538.
- Fligelstone, S. (1993). Some Reflections on the Question of Teaching, from a Corpus Linguistics Perspective. *ICAME Journal*, 17, 97-109.
- Fløttum, K. (2004). La présence de l'auteur dans les articles scientifiques : étude des pronoms *je*, *nous* et *on*. In A. Auchlin, E. Roulet & J.-M. Adam (Éd.), *Structures et discours. Mélanges offerts à Eddy Roulet*. Québec, Nota bene, 404-414.
- Fløttum, K., Dahl, T., & Kinn, T. (2006). *Academic Voices across languages and disciplines*. Amsterdam/Philadelphia : John Benjamins.
- Flowerdew, L. (2015). Learner corpora and language for academic and specific purposes. In S. Granger, G. Gilquin, & F. Meunier (Éd.), *The Cambridge Handbook of Learner Corpus Research*. Cambridge : Cambridge University Press, 465-484.
- Forsberg, F. (2006). *Le langage préfabriqué en français parlé L2. Étude acquisitionnelle et comparative* (Thèse de doctorat). Université de Stockholm : Cahiers de la recherche 34.
- Forsberg, F. (2010). Using conventional sequences in L2 French. *International Review of Applied Linguistics in Language Teaching*, 48(1), 25-51.
- François, J., & Sénéchal, M. (2006). Le sémantisme propre des cadres prédicatifs et la polysémie des verbes de production de parole. In D. Legallois & J. François (Éd.), *Autour des grammaires de construction et de patterns*. Université de Caen, 49-69.

---

Consulté le 23 octobre 2016 : <http://www.crisco.unicaen.fr/IMG/pdf/cahier21.pdf>

- Galligani, S. (2003). Réflexion autour du concept d'interlangue pour décrire des variétés non natives avancées en français. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 141- 152.
- Garcia-Debanc, C. (1990). *L'élève et la production d'écrits*. Metz. Centre d'Analyse Syntaxique de l'Université de Metz, Collection Didactique des textes.
- Gilquin, G. (2012). Lexical infelicity in English causative constructions. Comparing native and learner collocations. In J. Leino (Éd.), *Analytical causatives. Causatives. From 'give' and 'come' to 'let' and 'make'*. Lincom Europa : München, 41- 63.
- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora : The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6(4), 319- 335.
- Gilquin, G., & Paquot, M. (2008). Too chatty : Learner academic writing and register variation. *English Text Construction*, 1(1), 41- 61.
- Giuliani, D., & Hannachi, R. (2010). Linguistique de corpus et didactique du F.L.E. Une exploitation du corpus IntUne. *Cahiers de praxématique*, (54- 55), 145- 160.
- Gledhill, C. (2011). The 'lexicogrammar' approach to analysing phraseology and collocation in ESP texts. *ASp. La Revue Du GERAS*, (59), 5- 23.
- Goldberg, A. E. (2003). Constructions : a new theoretical approach to language, 7(5), 219- 224.
- Granger, S. (2002). A Bird's-eye View of Computer Learner Corpus Research. In S. Granger, J. Hung, & S. Petch-Tyson (Éd.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Benjamins : Amsterdam & Philadelphia, 3-33.
- Granger, S. (2007). Corpus d'apprenants, annotation d'erreurs et ALAO : une synergie prometteuse. *Cahiers de lexicologie*, 91(2), 117-132.
- Granger, S. (2008). Learner Corpora in Foreign Language Education. In N. Van Deusen-Scholl & N.H. Hornberger (Éd.), *Encyclopedia of Language and Education*, 4, *Second and Foreign Language Education*. US : Springer, 337- 351.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching : A critical evaluation. In K. Aijmer (Éd.), *Studies in Corpus Linguistics* (33). Amsterdam : John Benjamins Publishing Company, 13- 32. Consulté le 23 septembre 2016 : <https://benjamins.com/#catalog/books/scl.33.04gra/details>
- Granger, S. (2012). How to use foreign and second language learner corpora. In A. Mackey & S. M. Gass (Éd.), *Research methods in second language acquisition : a practical guide*. Oxford : Wiley-Blackwell, 7- 29.
- Granger, S., Gilquin, G., & Meunier, F. (Éd.). (2015). *The Cambridge Handbook of Learner Corpus Research*. Cambridge : Cambridge University Press.

- 
- Granger, S., & Monfort, G. (1994). La description de la compétence lexicale en langue étrangère : perspectives méthodologiques. *Acquisition et interaction en langue étrangère*, (3), 55-75.
- Granger, S., & Paquot, M. (2009a). In search of General Academic English : a corpus-driven study. In K. Katsampoxaki-Hodgetts (Éd.), *Proceedings of the International Conference on L.S.P. : Options and practices of L.S.P. practitioners*. Heraklion : University of Crete Publications, 94-108.
- Granger, S., & Paquot, M. (2009b). Lexical verbs in academic discourse : a corpus-driven study of learner use. In M. Charles, D. Pecorari, & S. Hunston (Éd.), *Academic writing : at the interface of corpus and discourse*. London : Continuum, 193-214.
- Granger, S., & Paquot, M. (2010). Customising a general EAP dictionary to meet learner needs. *Proceedings of ELEX2009 : ELexicography in the 21st century : New Challenges, new applications*, 87-96.
- Granger, S., & Paquot, M. (2015). Electronic lexicography goes local : Design and structures of a needs-driven online academic writing aid. *Lexicographica - International Annual for Lexicography/Internationales Jahrbuch für Lexikographie*, 31(1), 118-141.
- Granger, S., Vandeventer, A., & Hamel, M.-J. (2001). Analyse de corpus d'apprenants pour l'ELAO basé sur le TAL. *Traitement automatique des langues*, 42(2), 609-621.
- Gross, M. (1975). *Méthodes en syntaxe : régime des constructions complétives*. Paris : Hermann.
- Gross, M. (1986). *Grammaire transformationnelle du français : Syntaxe du verbe*. Paris : Cantilène.
- Grossmann, F. (2012). Pourquoi et comment cela change ? Standardisation et variation dans le champ des discours scientifiques. *Pratiques*, 153/154, 141-160.
- Grossmann, F. (2014a). De quelques routines phraséologiques liées aux verbes parenthétiques dans les genres scientifiques. In *SHS Web of Conferences 8*. EDP Sciences, Institut de Linguistique française, 759-770. Consulté le 17 mars 2015 : [http://www.shs-conferences.org/articles/shsconf/pdf/2014/05/shsconf\\_cmlf14\\_01306.pdf](http://www.shs-conferences.org/articles/shsconf/pdf/2014/05/shsconf_cmlf14_01306.pdf)
- Grossmann, F. (2014b). Les verbes de constat dans l'écrit scientifique. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presses Universitaires de Rennes, 85-100.
- Grossmann, F., & Tutin, A. (2010a). 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)*. Berlin, New York : De Gruyter Mouton, 279-307.
- Grossmann, F., & Tutin, A. (2010b). Les marqueurs verbaux de constat : un lieu de

- 
- dialogisme dans l'écrit scientifique. *Actes Du Colloque Dialogisme : Langue, Discours*. Consulté le 23 octobre 2016 : [http://www.univ-montp3.fr/praxiling/IMG/pdf\\_GrossmannTutin.pdf](http://www.univ-montp3.fr/praxiling/IMG/pdf_GrossmannTutin.pdf)
- Grossmann, F., Tutin, A., & Garcia Da Silva, P. (2009). Filiation et transfert d'objets scientifiques dans les écrits de recherche. *Pratiques*, 143/144, 187-202.
- Hanks, P. (2000). Contributions of Lexicograph and Corpus Linguistics to a Theory of Language Performance. *Euralex Proceedings*, (1), 3-13.
- Hanks, P. (2008). Lexical Patterns from Hornby to Hunston and beyond. In : E. Bernal & J. DeCesaris (Éd.), *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)*. Barcelona : Iula, 89-129.
- Hanks, P. (2013). *Lexical Analysis : Norms and Exploitations*. MIT Press.
- Hanks, P., & Pustejovsky, J. (2005). A Pattern Dictionary for Natural Language Processing. *Revue française de linguistique appliquée*, X(2), 63-82.
- Hartwell, L. M., & Jacques, M.-P. (2014). Authorial Presence in French and English : « Pronoun + Verb » Patterns in Biology and Medicine Research Articles. *Discours*, (varia 15). Consulté le 24 juin 2016 à l'adresse <http://discours.revues.org/8941>
- Hatier, S. (2013). Extraction des mots simples du lexique scientifique transdisciplinaire dans les écrits de sciences humaines : une première expérimentation. In *Actes de la 15e Rencontres des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL'2013)*. Les Sables d'Olonne, France, 138-149.
- Hatier, S. (2016). *Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche outillée sur un corpus d'articles de recherche en SHS* (Thèse de doctorat). Université Grenoble Alpes, Grenoble. Consulté le 13 février 2017 à l'adresse <http://www.theses.fr/s94109>.
- Hatier, S., Augustyn, M., Tran, H., Yan, R., Tutin, A., & Jacques, M.-P. (2016). French cross-disciplinary scientific lexicon extraction and linguistic analysis. In T. Margalitadze & G. Meladze (Éd.), *Proceedings Of The Xvii Euralex International Congress Lexicography & Linguistic Diversity*. Présenté à XVII EURALEX International congress, Tbilisi : Ivane Javakhishvili Tbilisi State University, 355-365.
- Hatier, S., Tutin, A., Jacques, M.-P., Jacquy, E., & Kister, L. (2014). Catégorisation sémantique des noms simples du lexique scientifique transdisciplinaire. Présenté à *Congrès ACFAS, colloque 330 : Étude de lexiques à vocation particulière : approches théoriques, méthodologiques, pédagogiques et multidisciplinaires*, 14 mai, Montréal. Consulté le 23 mai 2015 : <http://www.acfas.ca/evenements/congres/programme/82/300/330/c>
- Hatier, S., & Yan, R. (2017). Analyse contrastive des patrons verbaux dans l'écrit scientifique entre scripteurs étudiants et experts. *CORELA*, HS-21. Consulté le 20 mars 2017 à l'adresse <http://corela.revues.org/4879>



- 
- Heiden, S., Magué, J.-P., & Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In I. C. Sergio Bolasco (Éd.), *Proceedings of 10th International Conference on the Statistical Analysis of Textual Data*. Edizioni Universitarie di Lettere Economia Diritto, 2, Roma, Italy, 1021- 1032.
- Henderson, A. (2014). Le positionnement à travers la mention de l’auteur dans les écrits universitaires anglais d’étudiants français de premier cycle : implications pédagogiques. In A. Tutin & F. Grossmann (Éd.), *L’écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presses Universitaires de Rennes, 145-162.
- Hinkel, E. (2004). *Teaching Academic ESL Writing : Practical Techniques in Vocabulary and Grammar*. Mahwah, New Jersey & London : Lawrence Erlbaum Associates.
- Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford : Oxford University Press.
- Hou, H. (2008). *Erreurs linguistiques dans la production écrite en français chez les étudiants chinois* (Thèse de doctorat). Université des langues étrangères de Shanghai.
- Hunston, S., & Francis, G. (2000). *Pattern Grammar : a corpus-driven approach to the lexical grammar of English*. Amsterdam/Philadelphia : John Benjamins.
- Hyland, K. (2000). *Disciplinary Discourses : Social Interactions in Academic Writing*. Harlow, UK : Pearson Education.
- Hyland, K. (2004). Disciplinary Interactions : Metadiscourse in L2 Postgraduate Writing. *Journal of Second Language Writing*, 16, 133- 151.
- Hyland, K. (2005a). A convincing argument : corpus analysis and academic persuasion. In U. Connor & T. Upton (Éd.), *Discourse in the Professions : Perspectives from Corpus Linguistics*. Amsterdam : Benjamins, 87- 114.
- Hyland, K. (2005b). *Metadiscourse. Exploring Interaction in Writing*. Londres, New York, Continuum.
- Hyland, K. (2008a). academic clusters : text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1).
- Hyland, K. (2008b). As can be seen : Lexical bundles and disciplinary variation. *English for Specific Purposes*, (27), 4- 21.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students’ writing. *Journal of Second Language Writing*, 6(2), 183- 205.
- Jacques, M.-P. (2005). Pourquoi une linguistique de corpus ? In G. Williams (Éd.), *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes, 21-30.
- Jezeq, E., & Frontini, F. (2010). From Pattern Dictionary to Patternbank. *A Way with Words : Recent Advances in Lexical Theory and Analysis*, 215- 239.

- 
- Johns, T. (1991). Should you be persuaded : Two examples of data-driven learning. *Classroom concordancing. English language research journal*, 4, 1- 16.
- Kocourek, R. (1991). *La langue française de la technique et de la science : vers une linguistique de la langue savante*. Wiesbaden : O. Brandstetter.
- Kübler, N. (2014). Mettre en œuvre la linguistique de corpus à l'université : vers une compétence utile pour l'enseignement/apprentissage des langues ? *Recherches en didactique des langues et des cultures : Les Cahiers de l'Acedle*, 11(1), 37- 77.
- Lafont-Terranova, J., & Niwese, M. (2012). Acculturation à l'écriture de recherche et formation à la didactique de l'écriture. *Pratiques*, (153/154), 115- 128.
- Legallois, D. (2005). Du bon usage des expressions idiomatiques dans l'argumentation de deux modèles anglo-saxons : la Grammaire de Construction et la grammaire des Patterns. *Les Cahiers de l'Institut de Linguistique de Louvain*, 31(2-4), 109- 127.
- Legallois, D. (2006). La phraséologie dans la linguistique contextualiste. In D. Legallois & J. François (Éd.), *Autour des grammaires de construction et de patterns. Cahiers du CRISCO*, (21). Université de Caen, 28-48. Consulté le 12 janvier 2015 : <http://www.crisco.unicaen.fr/IMG/pdf/cahier21.pdf>
- Legallois, D., & Grea, P. (2007). L'objectif de cet article est de... Construction spécifique et grammaire phraséologique. *Cahiers de praxématique*, (46), 161- 186.
- Legallois, D., & Tutin, A. (2013). Présentation : Vers une extension du domaine de la phraséologie. *Langages*, 189(1), 3- 25.
- Lewis, M. (1993). *The lexical approach*. Hove : Language Teaching Publications.
- Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Éd.), *Second language vocabulary acquisition : A rationale for pedagogy*. Cambridge University Press, 255- 270.
- L'haire, S., & Vandeventer-Faltin, A. (2003). Diagnostic d'erreurs dans le projet FreeText. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, 6(2), 21-37.
- Liddicoat, A. (2008). *Discourse, genre and rhetoric : The French verb in research writing in science and technology*. Munich, Lincom.
- Lindsay, D., & Poindron, P. (2011). *Guide de rédaction scientifique : l'hypothèse, clé de voûte de l'article scientifique*. Versailles : Quae.
- Lorenz, G. (1999). Learning to cohere: causal links in native vs. non-native argumentative writing. In W. Bublitz, U. Lenk, & E. Ventola (Éd.), *Coherence in Spoken and Written Discourse. How to create it and how to describe it*. Amsterdam & Philadelphia : John Benjamins Publishing Company, 55- 75.
- Luste-Chaâ, O. (2008, 2009). *Les acquisitions lexicales en français langue seconde: conceptions et applications*. Université Paul Verlaine, Metz.

- 
- Luste-Chaâ, O. (2010). L'erreur lexicale dans l'analyse des productions écrites en FLE. *Pratiques*, 145- 146, 197- 210.
- Luzón, M. J. (2009). The use of we in a learner corpus of reports written by EFL Engineering students. *Journal of English for Academic Purposes*, 8(3), 192- 206.
- Mangenot, F., & Phoungsub, M. (2010). Un dispositif d'aide à la rédaction par incitations et socialisation. *Revue française de linguistique appliquée*, XV(2), 101- 119.
- Mangiante, J.-M., & Parpette, C. (2010). Faire des études supérieures en langue française. *Le français dans le Monde, Recherches & Applications*, (47), 53-63.
- Mangiante, J.-M., & Parpette, C. (2011). *Le français sur objectif universitaire*. Grenoble : Presses universitaires de Grenoble.
- Milton, J. C., & Chowdhury, N. (1994). Tagging the interlanguage of Chinese learners of English. *Proceedings joint seminar on corpus linguistics and lexicology*, 127- 143.
- Montagnon, G. (2007). *Développement d'une séquence didactique pour l'utilisation des collocations dans les écrits universitaires* (Mémoire de Master 1). Université Stendhal-Grenoble 3, Grenoble.
- Mroué, M. (2014). *Écrit de recherche universitaire : Éléments pour une sensibilisation au positionnement scientifique à travers la phraséologie transdisciplinaire* (Thèse de doctorat). Université de Grenoble Alpes, Grenoble.
- Nazar, R., & Renau, I. (2016). A quantitative analysis of the semantics of verb-argument structures. In S. Torner and E. Bernal (Éd.), *Collocations and other lexical combinations in Spanish. Theoretical and Applied approaches*. New York : Routledge, 92-109.
- Neff, J., Ballesteros, F., Dafouz, E., Martínez, F., & Rica, J. P. (2004). The expression of writer stance in native and non-native argumentative texts. In R. Facchinetti & F. Palmer (Éd.), *English Modality in Perspective*, 141- 161.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. McH. Sinclair (Éd.), *How to Use Corpora in Language Teaching*, 125- 152.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam : Benjamins.
- Noble, W. (2006). *An Analysis of Argument in Learner Academic Writing in a Second Language* (Ph.D. dissertation). Sydney : Macquarie University.
- Oksefjell Ebeling, S., & Hasselgård, H. (2015). Learner corpora and phraseology. In S. Granger, G. Gilquin, & F. Meunier (Éd.), *The Cambridge Handbook of Learner Corpus Research*. Cambridge : Cambridge University Press, 207- 229.
- O'Sullivan, I. (2010). Using corpora to enhance learners' academic writing skills in French. *Revue française de linguistique appliquée*, XV, 21- 35.
- O'Sullivan, I., & Chambers, A. (2006). *Learners' Writing Skills in French : Corpus*

- 
- Consultation and Learner Evaluation. *Journal of Second Language Writing*, 15(1), 49- 68.
- Paquot, M. (2010). *Academic Vocabulary in Learner Writing : From Extraction to Analysis*. Continuum International Publishing Group.
- Paquot, M. (2012). The LEAD dictionary-cum-writing aid : an integrated dictionary and corpus tool. In S. Granger & M. Paquot (Éd.), *Electronic Lexicography*. Oxford University Press, 163- 185.
- 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. (2005). Les apports possibles de la phraséologie à la didactique des langues étrangères. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, 8(1), 109-122.
- Pecman, M. (2007). Approche onomasiologique de la langue scientifique générale. *Revue française de linguistique appliquée*, XII(2), 79- 96.
- Pecman, M. (2012). Etude lexicographique et discursive des collocations en vue de leur intégration dans une base de données terminologiques. *The Journal of specialised translation (JoSTrans)*, 18, 113-138.
- Pecman, M., & Kübler, N. (2011). ARTES : an online lexical database for research and teaching in specialized translation and communication. *Proceedings from International Workshop on Lexical Resources (WoLeR) 2011*, 87- 93.
- Phal, A. (1970). *Vocabulaire général d'orientation scientifique (V.G.O.S.) - Part du lexique commun dans l'expression scientifique*. Paris : Didier.
- Phoungsub, M. (2012). Aides informatiques par incitations face aux problèmes de production écrite des apprenants thaïlandais : Quel impact sur leur processus d'écriture et leur dimension psychoaffective. *Procedia - Social and Behavioral Sciences*, 34, 181- 187.
- Pollet, M.-C. (2014). *L'écrit scientifique à l'aune des littéracies universitaires : approches théoriques et pratiques*. Namur : Presses universitaires de Namur.
- Pollet, M.-C., & Piette, V. (2002). Citations, reformulations du discours d'autrui : une clé pour enseigner l'écriture de recherche ? *Spirale*, (29), 165- 179.
- Popescu, O., Palmer, M., & Hanks, P. (2014). Mapping CPA Patterns onto OntoNotes Senses. Présenté à *LREC 2014, Ninth International Conference on Language Resources and Evaluation*, Reykjavik, Islande. Consulté le 02 juin 2016 à l'adresse [http://www.lrec-conf.org/proceedings/lrec2014/pdf/636\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/636_Paper.pdf)
- Poudat, C. (2006). *Études contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres* (Thèse de doctorat). Université d'Orléans, Orléans.
- Rastier, F. (2002). Enjeux épistémologique de la linguistique de corpus. In G. Williams

- 
- (Éd.). *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes, 31-45.
- Renau, I., & Battaner, P. (2012). Using CPA to represent Spanish pronominal verbs in learners' dictionary. In R.V. Fjeld, & J. M. Torjusen, (Éd.), *Proceedings of the XV EURALEX International Congress*. Oslo : University of Oslo, 350- 361.
- Reuter, Y. (1998). De quelques obstacles à l'écriture de recherche. *Lidil*, (17), 11- 23.
- Reuter, Y. (2004). Analyser les problèmes de l'écriture de recherche en formation. *Pratiques*, (121- 122), 9- 27.
- Reutner, U. (2010). De nobis ipsis silemus ? Les marques de personne dans l'article scientifique. *Lidil*, (41), 79- 102.
- Riegel, M., Pellat, J.-Ch., & Rioul, R. (1994, rev. 2009). *Grammaire méthodique du français*. Paris : P.U.F.
- Rinck, F. (2006). Écrire au nom de la science et de sa discipline : la figure de l'auteur dans l'article en sciences humaines. *Sciences de la Société*, 67, 95- 112.
- Rinck, F. (2010a). L'analyse linguistique des enjeux de connaissance dans le discours scientifique. *Revue d'anthropologie des connaissances*, 4(3), 427- 450.
- Rinck, F. (2010b). Les « nouveaux entrants » dans le champ scientifique : analyse des spécificités des articles de doctorants. In J.-M Defays, A. Englebert (Éd.), *Les discours universitaires*, tome 1 et 2. Paris : L'Harmattan, 99-110.
- Rinck, F. (2011). Former à (et par) l'écrit de recherche. Quels enjeux, quelles exigences ? *Le français aujourd'hui*, 3(174), 79- 89.
- Rinck, F., Boch, F., & Grossmann, F. (2007). Quelques lieux de variation du positionnement énonciatif dans l'article de recherche. In P. Lambert, A. Millet, M. Rispail, & C. Trimaille (Éd.), *Variations au coeur et aux marges de la sociolinguistique : mélanges offerts à Jacqueline Billiez*. Paris : L'Harmattan, 285-296.
- Rinck, F., & Jacques, M.-P. (2015). Corpus de littéracie avancée : structuration et méta-données. Présenté à *Journée d'études : Analyser informatiquement des grands corpus d'écrits scolaires. Problèmes de transcription, d'annotation et de traitement*. Sorbonne Nouvelle Paris 3 (18 mars 2015). Consulté le 25 avril 2016 : <http://www.univ-paris3.fr/activites-ecriscol-300518.kjsp?RH=1416243740283>
- Römer, U. (2010). Using general and specialized corpora in English language teaching : Past, present and future. In M. Carmen Campoy, B. Belles-Fortunato, & G.-V. Maria Lluisa (Éd.), *Corpus-based Approaches to English Language Teaching*. London : Continuum, 18- 35.
- Römer, U., & Wulff, S. (2010). Applying corpus methods to written academic texts : Explorations of MICUSP. *Journal of Writing Research*, 2(2), 99.
- Sánchez Cárdenas, B. (2010). *Paramètres linguistiques pour la conception d'un dictionnaire électronique bilingue (français-espagnol) destiné à la traduction : le*

---

*cas des verbes de comptage* (Thèse de doctorat). Université de Strasbourg.

- Schutz, N. (2012). *High-frequency verbs in EAP : towards a Corpus Patterns Analysis*. Présenté à *International Computer Archive of Modern and Medieval English (ICAME) 33 : Corpora at the centre and crossroads of English linguistics*. University of Leuven (30 May-3 June 2012). Consulté le 13 janvier 2015 : [https://www.researchgate.net/publication/269093698\\_High-frequency\\_verbs\\_in\\_EAP\\_towards\\_a\\_Corpus\\_Pattern\\_Analysis](https://www.researchgate.net/publication/269093698_High-frequency_verbs_in_EAP_towards_a_Corpus_Pattern_Analysis)
- Siepmann, D. (2007). Les marqueurs de discours polylexicaux en français scientifique. *Revue française de linguistique appliquée*, XII(2), 123- 136.
- Sinclair, J. (Éd.). (1987). *Looking Up : an account of the COBUILD Project in Lexical Computing*. London : Collins.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford : Oxford University Press.
- Sun, Y.-C. (2007). Learner perceptions of a concordancing tool for academic writing. *Computer Assisted Language Learning*, 20(4), 323- 343.
- Swales, J. M. (1990). *Genre analysis : English in academic and research settings*. Cambridge : Cambridge University Press.
- Tran, T. T. H. (2009). *Séquence pédagogique pour l'enseignement et l'apprentissage des collocations de démarcation* (Mémoire de Master 1). Grenoble : Université Grenoble Alpes.
- Tran, T. T. H. (2014). *Description de la phraséologie transdisciplinaire scientifique et réflexions didactiques pour l'enseignement à des étudiants non-natifs. Application aux marqueurs discursifs* (Thèse de doctorat). Grenoble : Université Grenoble Alpes.
- Tran, T. T. H., & Falaise, A. (2016). *Ergonomie des dictionnaires pour l'aide à la rédaction, état de l'art et propositions*. Présenté à *Journée « Work in progress » : autour des discours scientifiques*, Grenoble (juin 2016). Consulté le 23 mars 2016 : <https://pro.aiakide.net/publis/2016WIP-tran-falaise-slides.pdf>
- Tuchais, S. (2013). Les verbes d'opinion et la modalisation en français : état des lieux et perspectives. *Bulletin of the Faculty of Foreign Studies*, (47), 153- 172.
- Tutin, A. (2005). Le dictionnaire de collocations est-il indispensable ? *Revue française de linguistique appliquée*, X(2), 31- 48.
- Tutin, A. (2007a). 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. (2007b). Modélisation linguistique et annotation des collocations : une application au lexique transdisciplinaire des écrits scientifiques. In S. Koeva, D. Maurel, & M. Silberztein (Éd.), *Formaliser les langues avec l'ordinateur : de INTEX à Nooj*, 8. Besançon : Presses Universitaires de Franche Comté, 189-215.

- 
- Tutin, A. (2010a). Dans cet article, nous souhaitons montrer que...Lexique verbal et positionnement de l'auteur dans les articles en sciences humaines. *Lidil*, (41), 15-40.
- Tutin, A. (2010b). Evaluative adjectives in academic writing in the humanities and social sciences. In R. Lores-Sanz, P. Mur-Duenas, & E. Lafuente-Millan (Éd.), *Constructing Inerpersonality: Multiple Perspectives on Written Academic Genres*. Cambridge : Cambridge Scholars Publishing, 219-240.
- Tutin, A. (2010c). 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*. Louvain : Presses de l'Université de Louvain, 303-312.
- Tutin, A. (2014). 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*. Rennes : Presse universitaire de Rennes, 27-44.
- Tutin, A., & Falaise, A. (2013). Multiword expressions in scientific discourse : a corpus-driven database. In *Actes de eLex 2013*, Talinn, Estonia. Consulté le 16 avril 2016 à l'adresse <https://pro.aiakide.net/v2/publis/2013ElexPaper.pdf>
- Tutin, A., & Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue française de linguistique appliquée, Acquisition des langues : nouvelles orientations*, VII(1), 7-25.
- Tutin, A., & Grossmann, F. (2014). *L'Écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presse universitaire de Rennes.
- Tutin, A., Grossmann, F., Falaise, A., & Kraif, O. (2009). Autour du projet Scientext : étude des marques linguistiques du positionnement de l'auteur dans les écrits scientifiques. In *Actes des 6<sup>e</sup> Journées Internationales de la Linguistique de Corpus*. Lorient (10-12 septembre 2009), 333-349.
- Verlinde, S., Selva, T., & Binon, J. (2009). La base lexicale du français : de la lexicographie d'apprentissage à l'environnement d'apprentissage. In M. Heinz (Éd.), *Le dictionnaire maître de langue. Lexicographe et didactique*. Berlin : Franck et Timme, 289-306.
- Willems, D. (1981). *Syntaxe, lexique et sémantique des constructions verbales*. Gent : Rijksuniversiteit.
- Willems, D. (1983). Syntax and semantics : on the search of constants in verbal polysemy. In H. Shirö & I. Kazuko (Éd.), *Proceedings of the XIIIth International Congress of Linguistics*. Tokyo, Gakushuin University (August 29-September 4, 1982), 425-430.
- Williams, G. (Éd.). (2005). *La linguistique de corpus*. Rennes : Presses universitaires de Rennes.
- Williams, G. (2008a). The Good Lord and his works : A corpus-based study of collocational resonance. In S. Granger & F. Meunier (Éd.), *Phraseology : an*

- 
- interdisciplinary perspective*. Amsterdam : Benjamins, 159- 173.
- Williams, G. (2008b). Verbs of Science and the Learner's Dictionary. In E. Bernal & J. DeCesaris (Éd.), *Actes du 13<sup>e</sup> Congrès International d'Euralex*. Barcelona, Universitat Pompeu Fabra, 797-806.
- Williams, G., & Millon, C. (2010). Going organic : Building an experimental bottom-up dictionary of verbs in science. *Actes du 14<sup>e</sup> Congrès International d'Euralex*. Leeuwarden : Pays Bas, 1251- 1257.
- Williams, G., & Millon, C. (2013). Les verbes et la science : construction d'un dictionnaire organique. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presses Universitaires de Rennes, 45- 66.
- Williams, G., Millon, C., & Alonso, A. (2011). Collocational networks and their application to an E-Advanced Learner's Dictionary of Verbs in Science (DicSci). In *Proceedings of eLex 2011*. Bled, Slovenie, 12-22. Consulté le 30 mai 2015 à l'adresse <http://www.trojina.si/elex2011/Vsebine/proceedings/eLex2011-2.pdf>
- Williams, G., Millon, C., & Alonso, A. (2012). Growing naturally : The DicSci Organic E-Advanced Learner's Dictionary of Verbs in Science. In *Proceedings of the 15th EURALEX International Congress*. Oslo, University of Oslo, 1008- 1013.
- Wray, A. (2002). *Formulaic sequences and the lexicon*. Cambridge : Cambridge University Press.
- Wulff, S., & Gries, S. T. (2011). Corpus-driven methods for assessing accuracy in learner production. In P. Robinson (Éd.), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*. Amsterdam : Benjamins, 61- 87.
- Yan, R. (2014). Modalisation des patrons lexico-syntaxiques dans le corpus Scientext : exemple des verbes d'opinion. Présenté à *Colloque international des Etudiants-Chercheurs en Didactique des Langues et Linguistique*. Grenoble (juin 2014). Grenoble : Presses Universitaires de Grenoble. Consulté le 12 décembre 2017 : <https://hal.archives-ouvertes.fr/hal-01252355/document>
- Yan, R., & Hatier, S. (2016). L'extraction et la modélisation de patrons lexico-syntaxiques pour leur enseignement en FLE : un exemple à partir du verbe montrer. *Linguistik online*, 78(4), 93- 112.
- Ying, X. (2013). *Approche cognitive en didactique des langues : analyse et interprétation d'erreurs écrites prototypiques en français langue étrangère par des apprenants chinois et remédiation* (Thèse de doctorat). Université de Bordeaux 3, Université de Wuhan.
- Yoon, H., & Alain, H. (2004). ESL student attitudes towards corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257- 283.



---

# ANNEXE

# ANNEXE 1

## Liste des verbes du LST

verbe	Fréquence Scientext-SHS	Ratio (Scientext-SHS/corpus de contraste)	verbe	Fréquence Scientext-SHS	Ratio (Scientext-SHS/corpus de contraste)	verbe	Fréquence Scientext-SHS	Ratio (Scientext-SHS/corpus de contraste)
permettre	5088	3.625	conduire	1475	1.988	distinguer	969	5.205
montrer	3053	2.343	correspondre	1469	9.200	imposer	953	2.489
présenter	2663	2.390	observer	1450	3.004	évoquer	947	2.340
agir	2594	2.126	lier	1412	4.767	situer	937	3.548
constituer	2492	7.326	développer	1389	5.207	effectuer	936	3.071
considérer	2480	5.214	construire	1323	3.494	citer	935	4.552
utiliser	2308	2.649	impliquer	1228	5.655	analyser	932	8.196
sembler	2302	1.507	associer	1221	8.347	viser	925	3.049
proposer	2193	2.404	concerner	1159	3.472	identifier	920	4.378
apparaître	2044	4.167	souligner	1146	3.275	offrir	919	1.214
devenir	2026	1.143	décrire	1105	5.119	indiquer	901	1.718
porter	1687	1.088	noter	1088	3.172	inscrire	895	3.198
définir	1630	13.655	supposer	1075	2.545	déterminer	884	5.323
représenter	1623	2.954	exprimer	1056	3.119	reconnaître	876	1.055
obtenir	1621	2.429	relever	1010	2.258	réduire	863	2.622
expliquer	1584	1.038	caractériser	1008	37.086	marquer	854	2.173
produire	1529	2.791	établir	1001	4.001	consister	839	6.070
réaliser	1486	2.422	étudier	977	3.350	paraître	839	0.906
exister	1482	2.309	mener	973	1.312	désigner	834	3.080

intégrer	827	6.037	signifier	651	2.140	juger	500	0.946
fonder	826	6.494	accorder	645	2.164	qualifier	498	2.676
contribuer	819	6.656	préciser	644	1.532	posséder	496	1.477
choisir	817	1.069	atteindre	643	1.073	comporter	494	3.003
favoriser	809	6.771	adopter	638	2.951	employer	493	2.130
organiser	798	1.186	consacrer	630	2.329	interpréter	484	4.482
renvoyer	797	4.069	retenir	626	1.403	résumer	483	3.335
limiter	779	4.207	confirmer	619	1.433	manifester	479	2.738
appuyer	771	2.452	estimer	614	1.061	affecter	475	3.970
participer	761	1.769	intervenir	609	1.716	conserver	472	1.843
révéler	759	1.981	constater	608	1.806	entraîner	466	1.177
interroger	754	1.879	appartenir	608	1.553	résulter	464	12.615
reposer	752	2.518	modifier	602	3.942	adapter	464	2.680
tendre	749	1.488	accompagner	601	1.233	soutenir	464	1.178
dépendre	746	3.827	appliquer	587	3.040	illustrer	455	6.045
intéresser	744	1.650	apporter	587	0.948	justifier	454	2.719
former	735	1.930	privilégier	579	7.507	pratiquer	450	2.470
disposer	728	2.006	concevoir	576	3.390	vérifier	450	1.133
fournir	720	2.847	composer	571	2.184	attribuer	448	3.959
placer	712	0.907	exercer	569	3.675	examiner	444	1.463
partager	710	1.693	mobiliser	557	4.431	avérer	437	5.230
augmenter	700	2.282	renforcer	550	2.812	aborder	433	2.915
traduire	692	4.521	témoigner	549	3.103	saisir	430	1.079
affirmer	684	1.315	opérer	542	4.028	transformer	429	1.449
envisager	681	2.845	amener	538	0.996	soumettre	428	2.580
opposer	673	2.604	maintenir	534	1.057	accroître	414	5.578
publier	673	2.140	demeurer	527	1.866	figurer	412	1.928
élever	671	1.672	suggérer	526	2.408	centrer	411	6.028
tenter	663	0.896	mesurer	512	2.962	contenir	407	1.499
évaluer	659	6.615	varier	511	8.893	compléter	406	3.490
traiter	658	1.659	contrôler	511	1.500	élaborer	396	6.996
introduire	657	5.252	comparer	506	4.076	insister	390	1.264
importer	654	1.055	convenir	500	2.264	structurer	389	20.421

calculer	388	5.021	provenir	314	2.659	généraliser	243	7.691
entretenir	385	2.551	diffuser	314	2.074	rassembler	243	0.950
admettre	383	1.208	respecter	312	1.040	générer	240	5.136
provoquer	379	1.176	baser	309	2.451	valider	238	4.563
confronter	377	3.709	engendrer	306	6.521	dessiner	237	1.405
conclure	377	1.488	rapprocher	303	1.549	recourir	233	5.906
dépasser	376	1.206	différencier	302	18.097	combiner	231	8.560
regrouper	367	3.332	formuler	300	4.505	recueillir	229	1.694
attacher	361	1.617	résoudre	300	1.393	émerger	228	3.051
exclure	360	3.134	satisfaire	294	1.680	revendiquer	228	2.712
tester	355	2.878	contraindre	293	2.528	attester	225	8.410
procéder	355	2.319	aboutir	290	3.198	différer	225	7.460
inclure	353	6.127	améliorer	290	1.649	élargir	221	2.914
fonctionner	353	1.404	résider	288	4.368	conférer	220	7.418
nommer	346	1.600	garantir	286	2.025	articuler	218	4.978
autoriser	344	1.435	signaler	284	1.746	assumer	217	1.876
appréhender	342	10.188	séparer	280	1.071	confondre	216	2.431
faciliter	342	3.573	influencer	273	6.683	corrélér	214	104.285
relier	339	2.966	rejeter	272	1.510	distribuer	209	1.819
classer	339	1.936	restreindre	271	8.362	exploiter	207	3.046
diminuer	335	3.208	refléter	271	4.830	répartir	207	3.024
évoluer	335	1.993	accéder	268	2.437	éclairer	206	1.250
étendre	334	1.600	remplacer	264	0.959	reproduire	202	2.425
mentionner	333	3.872	rechercher	262	1.417	émettre	199	1.129
nécessiter	321	4.851	expliciter	260	35.910	emprunter	198	1.282
dégager	320	1.090	inciter	257	2.488	énoncer	196	8.804
repérer	318	1.534	concentrer	256	1.275	explorer	190	2.210
induire	317	15.991	multiplier	255	1.795	localiser	189	2.620
référer	317	14.140	précéder	252	2.525	sélectionner	186	3.146
doter	317	2.865	critiquer	252	2.281	spécifier	178	27.203
discuter	317	1.009	prétendre	248	1.080	diviser	178	1.437
démontrer	315	2.991	inspirer	247	1.305	déduire	175	4.359
exposer	315	1.326	exiger	245	1.047	renouveler	175	1.811

consulter	174	1.113	reporter	127	1.519	objectiver	90	28.347
découler	172	8.938	nuancer	126	9.433	reconstruire	95	2.205
substituer	172	7.990	recenser	125	2.695	préconiser	94	2.233
focaliser	171	6.944	approfondir	124	5.931	coïncider	93	4.882
prolonger	171	1.899	alimenter	124	1.351	conjuguer	93	4.710
hériter	168	2.049	contester	124	1.226	englober	92	6.770
recouvrir	163	1.217	invoquer	123	2.307	objectiver	90	28.347
assimiler	162	5.378	reconstituer	121	2.566	affiner	88	5.612
ressortir	155	1.288	progresser	121	0.937	subsister	88	2.303
inverser	154	3.504	simplifier	120	3.703	retracer	87	2.616
efforcer	153	1.047	ajuster	120	2.927	redéfinir	86	10.577
instituer	152	8.724	décliner	119	2.091	amorcer	86	2.756
accentuer	151	3.173	débattre	117	1.069	converger	85	4.889
rattacher	147	5.886	collecter	115	3.081	équivaloir	85	3.024
assigner	146	4.884	corriger	115	1.591	esquisser	85	1.943
enrichir	146	2.606	sous-tendre	112	29.822	dissocier	83	10.208
guider	146	1.033	cerner	110	2.668	relativiser	83	4.523
motiver	144	2.037	questionner	108	1.764	hiérarchiser	82	19.253
compenser	138	2.211	succéder	107	0.917	symboliser	82	3.273
postuler	134	9.180	conformer	106	3.631	démarquer	80	5.395
conditionner	132	4.955	formaliser	104	12.436	affranchir	80	4.368
délimiter	131	7.742	contraster	103	6.634	raisonner	78	2.387
apparenter	131	5.088	véhiculer	101	5.538	médiatiser	77	7.393
conforter	131	3.502	masquer	101	1.737	advenir	74	2.435
isoler	131	1.697	ancrer	100	3.398	amplifier	74	2.096
heurter	131	1.036	atténuer	98	2.880	influer	73	5.082
négliger	130	2.509	reconstruire	95	2.205	occulter	73	3.002
détailler	130	2.156	préconiser	94	2.233	replacer	69	3.000
déboucher	130	1.797	coïncider	93	4.882	dispenser	69	1.056
restituer	127	4.085	conjuguer	93	4.710	prôner	66	1.705
revêtir	127	3.166	englober	92	6.770	contredire	56	1.720

## ANNEXE 2

### Classification sémantique des verbes du LST

Classe	Sous-classe	Verbes
<b>analyse</b> Analyse des informations : description des données, analyse des relations entre les variables, comparaison des résultats attendus, interprétation des écarts, analyse statistique, résultats et conclusions	<b># inférence</b> Faire sortir d'un fait ou d'un évènement la conséquence logique qu'elle contient implicitement. N[sujet] +V+ N[COD]/ de N que P	conclure_2, déduire
	<b>#comparaison</b> Action scientifique qui vise à rapprocher pour mettre en évidence des rapports de ressemblance ou de différence. N[sujet] +V+ N[COD] (à/de N2)/passif	apparenter, assimiler, comparer, confronter_1, opposer_1, rapprocher_1
	<b>#compréhension</b> Saisir par un acte précis de la pensée conceptuelle en élaborant dans son esprit la représentation nette d'une chose. N[sujet] +V+ N[COD]/que P	appréhender, réaliser_2, résoudre, saisir_1
	<b>#constat</b> Action scientifique qui vise à prendre connaissance de qqch au moyen d'observations scientifiques. N[sujet] +V+ N[COD]/ que P	constater, identifier, noter, observer, relever_2, repérer_1, repérer_2,
	<b>#démonstration</b> Action scientifique qui vise à mettre en évidence, prouver la vérité, la réalité, la valeur de qqch. N[sujet] +V+ N[COD]	démontrer_1, établir_4, montrer_2
	<b>#description</b> Action scientifique qui vise à représenter en détail (par écrit ou oralement), certains traits apparents d'une entité à l'aide d'un outil. N[sujet] +V+ N[COD]	caractériser_1, décrire_1, détailler, illustrer, présenter_1, représenter_1, retracer, spécifier_1, spécifier_2

	<b>#détermination</b> Action scientifique qui vise à délimiter des propriétés de qqch.	définir_1, déterminer_1
	<b>#évaluation quantitative</b> Déterminer, marquer (par un jugement favorable ou défavorable) une quantité. <sup>[L]</sup> <sub>[SEP]</sub> N[sujet] +V+ N[COD]	calculer_1, estimer_1, évaluer_1, recenser
	<b>#évaluation qualitative</b> Déterminer, marquer (par un jugement favorable ou défavorable) une qualité, la valeur de qqch, le résultat d'une action. N[sujet] +V+ N[COD]	mesurer_07, mesurer_08 calculer_2, estimer_3, évaluer_2, mesurer_1, mesurer_2
	<b>#examen</b> Action scientifique qui vise à étudier en détail, faire l'analyse de qqch pour en avoir une connaissance approfondie, juger de sa valeur. N[sujet] +V+ N[COD]/passif	aborder, analyser, approfondir, considérer_2, contrôler, critiquer, étudier, examiner, explorer_1, explorer_2, soumettre_1, tester, traiter_1, traiter_2, vérifier_1
	<b>#identification</b> Action scientifique qui vise à <sup>[L]</sup> <sub>[SEP]</sub> approcher en vue d'une saisie précise. L'objet désigne qqch d'abstrait ou une réalité considérée abstraitement, qui passe pour difficile à saisir. N[sujet] +V+ N[COD] (de N2/que P)	cerner, dégager, distinguer_2, rechercher
	<b>#interprétation</b> Donner un sens personnel, parmi d'autres possibles, à un acte, à un fait, dont l'explication n'apparaît pas de manière évidente. N[sujet] +V+ N[COD] N[sujet] +V	éclairer_1, expliciter_1, expliquer_1, interpréter, justifier_1, raisonner
	<b>#réalisation</b> Action scientifique qui vise à obtenir un résultat, mener à bien un projet par la mise en œuvre de moyens appropriés. N[sujet] +V+ N[COD]	atteindre_3, construire, élaborer, esquisser, établir_2, générer_1, instituer, publier, réaliser_1, reconstituer, reconstruire, reproduire_1, restituer
	<b>#réexamen</b> Reconsidérer qqch d'abstrait en y apportant des informations supplémentaires, des éclaircissements et des modifications. N[sujet] +V+ N[COD]	nuancer, préciser_1, redéfinir, relativiser_1

<b>ancrage_construction</b> (construction du modèle d'analyse : problématique, hypothèses, conceptualisation, définition, réflexion, références)	<b>#choix</b> Se décider, se déterminer entre deux ou plusieurs choses; opter pour un parti, pour une solution. N[sujet] +V+ N [COD] <sup>[1]</sup> <sub>[SEP]</sub> de V/passif	admettre_1, adopter, choisir_1, choisir_2, intégrer_2, privilégier, retenir_1, sélectionner
	<b>#conception</b> Processus par lequel on forme ou crée dans son esprit une idée, un projet, une œuvre. N[sujet] +V+ N[COD]/passif	concevoir_1, constituer_1, structu_2, structure_2, structure316, former_1, structure_1
	<b>#dénomination</b> Action scientifique qui vise à attribuer un nom à qqch. N[sujet] +V+ N[COD]	désigner, nommer, qualifier
	<b>#fondement</b> Action scientifique qui vise à prendre qqch pour base, pour principe fondamental et de démontrer une chose en justifiant la démonstration par ces principes (des faits, des théories, des preuves, une documentation) qui soutiennent l'hypothèse. N[sujet] +V+ sur/de/dans N[CO] / passif	ancrer, appuyer_1, appuyer_s'_2, baser_1, baser_se_2, établir_1, fonder, inspirer_s'_2, reposer_se_2
	<b>#hypothèse</b> poser qqch comme hypothèse servant de base à un raisonnement, à une argumentation dans une recherche. N[sujet] +V+ N[COD]/que P	postuler, supposer_1
	<b>#finalité</b> Avoir en vue et mettre en œuvre son action dans un but déterminé N[sujet] + à/de + V/N N[sujet] +V+ N	attacher_s'_2, consacrer_se_1, efforcer_se, employer_s'_1, envisager_2, proposer_se_2, tendre_1, tenter, viser_1
	<b>#référence</b> Action scientifique qui vise à se servir de qqch comme appui, repère ; faire référence à, idée de faire appel, recourir à qqch (différence avec #fondement) N[sujet] +V+ à N[COI]	référer_1, référer_se_2, reporter_se_1
	<b>#restriction</b> Déterminer avec précision les limites de qqch. Être V + à/de + V/N N[sujet] +V+ N	délimiter_1, limiter_1, limiter_se_2, restreindre
	<b>#thème</b>	centrer, concentrer_2, consulter, focaliser_se,



	Rassembler ses forces intellectuelles ou affectives et les porter sur un certain objet, à l'exclusion de tout autre ; y appliquer avec force (sa pensée, son attention, etc.). N[sujet] +V+ sur N[COI] <sub>SEP</sub> N[sujet] +V+ qn + sur+ N N[sujet] +V+ N[COD]	interroger_s'_1, questionner
<b>communication</b> Classe regroupant les verbes de communication et de transfert des idées (selon le type de message, le moyen et la manière de transmission, faisant ou non partie d'une interaction).	<b>#discussion</b> Acte communicatif par lequel on procède à une prise de parole dans une discussion pour donner son avis ; l'examen contradictoire d'une question, d'une affaire; en débattre. N[sujet] +V+ N[COD]/de N N[sujet] +V	débattre, discuter, intervenir_2
	<b>#faire savoir</b> Acte communicatif par lequel on rend manifeste par toutes sortes de signes (langage écrit, oral, etc.), de façon volontaire ou non, ce que l'on pense. N[sujet] +V+ N[COD]/que P	diffuser, exposer, exprimer_1, indiquer_1, marquer_1, médiatiser, mentionner_2, préciser_2, révéler_1, signaler_1, véhiculer
	<b>#formulation</b> Acte communicatif par lequel on énonce avec la concision et la netteté d'une formule. N[sujet] +V+ N[COD]	émettre, énoncer, évoquer_2, formuler, invoquer, résumer_1
	<b>#mention</b> Acte communicatif par lequel on fait remarquer qqch par la parole ou par un écrit en signalant, sans s'étendre. N[sujet] +V+N[COD]	citer_1, citer_2, mentionner_1
	<b>#proposition</b> Acte communicatif par lequel on présente qqch d'abstrait (objet=analyse, modèle, projet, solution, définition, approche, etc.) à la réflexion, à l'examen ou à l'évaluation de qqch. N[sujet] +V+ N[COD] (à N2 de V/que P)	proposer_1, soumettre_2, suggérer_1
	<b>#mise en relief</b> Action d'attacher et de donner une importance particulière à qqch.	insister, souligner
<b>état</b> Classe des verbes renvoyant à l'état au sens général : une manière d'être, une situation ou condition	<b>attribution</b> L'état d'une entité qui fournit qqch, contribue à qqch. N[sujet] +V+ N[COD]	apporter_2, conférer, consacrer_2, doter_2, fournir_2, offrir_2

statique.	<p><b>composition</b> L'état d'une entité qui constitue assemblage ou combinaison de plusieurs éléments ou parties ; dotée d'une certaine structure. N[sujet] +V+ (de,en) N[COD/COI]/passif</p>	composer_se_1, consister_3, former_2, structurer_2
	<p><b>représentation</b> L'état d'avoir la même valeur qualitative, d'avoir pour équivalence ou de constituer le trait distinctif ou dominant de qqch tout en le rendant reconnaissable. N[sujet] +V+ N[COD]/ dans/à N[sujet] +V+ N[COD]/passif</p>	caractériser_2, composer_2, consister_1, équivaloir, représenter_2, représenter_4, résider, signifier, symboliser
	<p><b>existence</b> L'état de se montrer dans sa réalité, de se trouver dans une certaine situation. N[sujet/impersonnel ] +V+ N[COD] N[sujet/ impersonnel] +V+ de N[COI]/que P</p>	agir_s'_2, apparaître_2, avérer_s'_2, constituer_2, demeurer_1, demeurer_2, exister_2
	<p><b>fondement</b> L'état d'être fondé sur qqch, d'avoir son essence, ses propriétés dans qqch. N[sujet] +V+ à/sur N[COI]</p>	consister_2, reposer_1
	<p><b>inclusion</b> L'état de comprendre, de réunir en soi par nature des éléments ou des caractères qualitatifs ou quantitatifs. N[sujet] +V+ N[COD]/de N</p>	admettre_3, comporter_2, contenir, disposer_2, inclure, introduire_2, partager_2, porter_1, posséder_2, présenter_2, rassembler_2, regrouper_3, reporter_2, revêtir
	<p><b>localisation</b> L'état de trouver sa place dans des limites spatiales déterminées. N[sujet] +V (+CC)</p>	figurer, précéder_2, situer_se_1
	<p><b>Manifestation</b> L'état de se manifester, d'émerger, venir à la conscience.</p>	affirmer_2, avérer_s'_1, dessiner_se_1, exister_1, imposer_s'_2, manifester_se, refléter_se_2, traduire
	<p><b>rapport</b> L'état d'avoir quelque rapport avec qqch, avoir pour objet qqch.</p>	affecter, concerner, intéresser_1, porter_3, renvoyer_2, viser_2

	N[sujet] +V+ N[COD]/ à/sur N	
<b>modalité</b> Expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé.	<b>#obligation</b>	contraindre_1, imposer_1, soumettre_se_3
	<b>#permission</b>	autoriser_1
	<b>nécessité</b> Évaluation de la nécessité.	convenir_2, exiger, importer_1, nécessiter
	<b>obligation</b>	importer_2
	<b>permission</b>	autoriser_2, permettre_1
	<b>probabilité</b> évaluation des chances de réalisation de la relation prédicative, la probabilité	paraître_1, paraître_2, sembler
	<b>restriction</b> Apporter des limites à une entité.	contraindre_2, délimiter_2, réduire_se_1
<b>observation</b> Sélection d'unités d'observation, collecte des données, recueil des informations, expérimentation, tests	<b>#classement</b> Action scientifique qui vise à distribuer (objet=données, objet d'étude) dans des catégories distinctes selon des critères définis. N[sujet] +V+ N[COD]	classer_1, hiérarchiser
	<b>#collecte</b> Action scientifique qui vise à rassembler toute sorte d'éléments concrets ou abstraits (Objet=généralement données scientifiques) en vue d'une action précise. N[sujet]+V+ N[COD] / passif	collecter, concentrer_1, confondre, englober, rassembler_1, recueillir, regrouper_1, regrouper_2
	<b>#inclusion</b> Action scientifique qui vise à introduire un élément (objet concret ou abstrait) dans un ensemble afin que, s'y incorporant, il forme un tout cohérent. N[sujet] +V+ N[COD]	inscrire_1, intégrer_1, introduire_1
	<b>#séparation</b> Action scientifique qui vise à distinguer toute sorte d'éléments concrets ou abstraits (Objet=généralement données scientifiques) en vue d'une action précise. N[sujet] +V+ N[COD]	dispenser, dissocier, distinguer_3, diviser, isoler, répartir_1, répartir_2, séparer_1
<b>processus_cognitif_psycho</b>	<b>#conformité</b>	accorder_s' 2, rapprocher_se_2, rejoindre

Classe regroupant les verbes renvoyant processus cognitifs et psychologiques ne relevant pas prioritairement de l'activité scientifique.	Se rendre proche par rapport à qqch, à qqn en retrouvant certains traits communs.	
	<b>#correspondance</b> Se mettre en accord avec d'autres individus/entités/circonstances mettant en évidence les ressemblances, la conformité. N[sujet] +V+ à N[COI]	adapter_s'_1, conformer
	<b>#démarcation</b> Se rendre différent par rapport à qqch, à qqn. N[sujet] +V+ de N[COI]	affranchir_1, démarquer_se, différencier_se_1
	<b>#opinion défavorable</b> Prise de position par rapport à l'objet, mise en discussion, en doute. N[sujet] +V+ N[COD]/à +N[COI]/que P	contester, exclure_1, négliger, opposer_s'_2, rejeter
	<b>#opinion favorable</b> <sup>[L]</sup> <sub>[SEP]</sub> Prise de position par rapport à l'objet, jugement positif sur la valeur, sur l'utilité, approbation. N[sujet] +V+ N[COD]/que P	favoriser, préconiser, prôner, soutenir_2
	<b>#opinion neutre</b> Considération générale sans jugement ou orientation particulière. N[sujet] +V+ N[COD]/que P	admettre_2, affirmer_1, assumer, concevoir_2, considérer_1, considérer_3, envisager_1, estimer_2, garantir_2, juger, prétendre, reconnaître, représenter_se_3, revendiquer, soutenir_1, supposer_2
	<b>#positionnement</b> Prise de conscience sur la position/la place vis-à-vis de qqch. N[sujet] +V (+CC)	placer_se_2
	<b>#processus psychologique</b> Processus psychologiques généraux.	convenir_1, intéresser_2, montrer_se_3, partager_1, permettre_se_2, présenter_se_3, respecter, satisfaire_1
<b>processus évolutif</b>	<b>#amélioration</b> Processus évolutif par lequel on rend meilleur, plus satisfaisant, on change en mieux qqch. N[sujet] +V+ N[COD]	affiner, améliorer, compenser, compléter, corriger, faciliter, objectiver, relever_1, renouveler, simplifier

	<p><b>#augmentation</b> Processus évolutif par lequel on rend plus considérable (en quantité, en durée, etc.) ou plus fort (en intensité), en ajoutant à ce qui existe qqch de même nature. N[sujet] +V+ N[COD]</p>	<p>accentuer, accroître, amplifier, augmenter_1, développer_1, élever_1, multiplier_1</p>
	<p><b>#baisse</b> Processus évolutif par lequel on rend moins considérable (en quantité, en durée, etc.) ou plus fort (en intensité), par retranchement partiel. N[sujet] +V+ N[COD]</p>	<p>atténuer, diminuer_1, réduire_2</p>
	<p><b>#changement</b> Processus évolutif par lequel on rend plus ou moins différent, on fait subir une modification, on remplace en mettant à la place qqch de différent mais de même nature ou fonction. N[sujet] +V+ N[COD]</p>	<p>alimenter, développer_3, enrichir, inverser, modifier_1, remplacer, substituer_2, varier_1</p>
	<p><b>#extension</b> Processus évolutif par lequel on augmente l'importance, la portée de qqch. N[sujet] +V+ N[COD]</p>	<p>élargir, étendre, généraliser</p>
	<p><b>amélioration</b> Processus évolutif par lequel qqch devient meilleur, plus satisfaisant, change en mieux. N[sujet] +V+ N[COD]</p>	<p>relativiser_2</p>
	<p><b>augmentation</b> Processus évolutif par lequel qqch devient plus considérable (en quantité, en durée, etc.) ou plus fort (en intensité). N[sujet] +V N[sujet] +V+ N[COD]</p>	<p>augmenter_2, dépasser, élever_s'_2, multiplier_2, progresser_1, renforcer</p>
	<p><b>baisse</b> Processus évolutif par lequel qqch devient moins considérable (en quantité, en durée, etc.) ou plus fort (en intensité). N[sujet] +V N[sujet] +V+ N[COD]</p>	<p>décliner, diminuer_2</p>
	<p><b>changement</b> Processus évolutif par lequel qqch devient plus ou moins différent, subit une modification ou change progressivement</p>	<p>devenir, évoluer, modifier_2, progresser_2, prolonger_se, substituer_se_1, transformer_se_1, transformer_2, varier_2</p>

	de position ou de nature. N[sujet] +V N[sujet] +V+attribut N[sujet] +V+ N[COD]/à N[sujet] +V (vers)	
<b>processus_humain</b> Classe regroupant les verbes renvoyant aux processus humains généraux . (il s'agit des processus qui ne relèvent ni du cognitif ni du scientifique)	<b>#action</b> Réaliser une action, une transformation ou une opération complexe N[sujet] +V (+CC de manière)	agir_1, comporter_se_3, procéder_3
	<b>#attribution</b> Processus par lequel on donne qqch, une propriété, à un élément. N[sujet] +V+ N[COD] à N2	accorder_1, apporter_1, assigner, attacher_1, attribuer, distribuer, doter_1, fournir_1, offrir_1, porter_2
	<b>#composition</b> Combiner les éléments d'un ensemble d'une certaine manière. N[sujet] +V+ N[COD]	organiser
	<b>#faire</b> Verbes supports (transitifs) désignant la réalisation de qqch. N[sujet] +V+ N[COD]	conduire_1, effectuer, exercer, mener_2, opérer_1, procéder_1
	<b>#localisation</b> Attribuer un emplacement à (une chose), mettre dans un ordre déterminé, dans un ensemble. N[sujet] +V+ N[COD] (+CC)	localiser, placer_1, remplacer, situer_2
	<b>#obtenir</b> Avoir à soi, disposer de (qqch), parvenir à se faire accorder, à se faire donner. N[sujet] +V+ N[COD]/à N/de N	accéder_1, atteindre_1, disposer_1, obtenir_1, obtenir_2, posséder_1, retenir_2, saisir_se_2
	<b>#usage</b> Faire usage de qqch, appliquer un procédé, une technique, faire agir un objet, une méthode afin d'obtenir un effet qui permette d'aboutir à une fin déterminée. N[sujet] +V+ N[COD]/à N	appliquer_1, appliquer_2, employer_2, emprunter, exploiter_1, mobiliser, pratiquer, recourir_1, recourir_2, utiliser
<b>processus_non humain</b> Classe regroupant les verbes renvoyant aux processus non humains qui ne sont pas statiques, qui expriment	<b>action</b> Remplir (dans un ensemble) une fonction déterminée. N[sujet] +V (+CC de manière)	comporter_se_1, fonctionner_1, fonctionner_2

une action.	<b>choix</b>	exclure_2
	<b>dissimulation</b> Présenter une apparence qui cache qqch, qui empêche que l'on se rende compte ou que l'on prenne conscience de qqch. N[sujet] +V+ N[COD]	masquer, occulter
	<b>événement</b> (En parlant d'un phénomène, d'un évènement, d'une chose) venir à l'existence, arriver, survenir, le plus souvent après une succession d'autres faits. N[sujet] +V	apparaître_3, intervenir_1, produire_se_2, reproduire_se_2
	<b>obtenir</b> Disposer de (qqch), parvenir à se faire accorder. N[sujet] +V+ N[COD]/à N	accéder_2, atteindre_2
	<b>révélation</b> Rendre manifeste, laisser apparaître clairement la vérité d'un fait, les caractéristiques d'une entité abstraite. N[sujet] +V+ N[COD]/que P	décrire_2, démontrer_2, dessiner_2, éclairer_2, expliciter_2, exprimer_2, indiquer_2, marquer_2, montrer_1, refléter_1, révéler_se_2, signaler_2, témoigner
	<b>usage</b> Faire usage de qqch, appliquer un procédé, une technique, faire agir un objet, une méthode afin d'obtenir un effet qui permette d'aboutir à une fin déterminée. N[sujet] +V+ N[COD]	exploiter_2
	<b>attestation</b> Constituer la preuve, la justification de la vérité, de la validité de qqch. N[sujet] +V+ N[COD]/que P	attester, confirmer, conforter, garantir_1, justifier_2, valider, vérifier_2
	<b>contestation</b> Contester la vérité, la validité de qqch. N[sujet] +V+N[COD]	interroger_2
<b>relation</b> Classe regroupant les verbes renvoyant aux différents types de relation entre entités ou exprimant une mise en relation effectuée selon différents critères.	<b>#appartenance</b> Relation renvoyant pour un individu au fait d'appartenir à un milieu, une collectivité. N[sujet] +V+ à N[COI]	appartenir_1, participer_2

	<p><b>#association</b> Action de mettre en relation, réunir en un tout des éléments divers qui se complètent et produisent un effet unique. N[sujet] +V+ N[COD] + à/avec N2</p>	accompagner_1, articuler, associer, combiner_1, conjuguer, corréler, lier, rattacher, relier
	<p><b>#cause_objectif</b> Relation causale avec indication du point d'aboutissement. N[sujet] +V+ N[COD] + à N[COI]/ Vinf</p>	amener, inciter, renvoyer_1
	<p><b>#influence</b> Relation renvoyant au fait de produire un effet suscitant des modifications, manière d'agir sur qqn ou qqch. N[sujet] +V+ N[COD]</p>	influencer, influer_1
	<p><b>#opposition</b> Relation caractérisant qqn par le rapprochement avec d'autres, mettant en évidence leurs différences. N[sujet] +V+ à N[COI]</p>	confronter_se_2, heurter_se
	<p><b>appartenance</b> Relation caractérisant un état d'une entité par rapport à un ensemble dont elle fait partie. N[sujet] +V+ à/de/N[COI]</p>	appartenir_2, inscrire_s'_2, participer_1, relever_3, ressortir_2
	<p><b>association</b> Relation caractérisant un état d'une entité mise ensemble, réunie avec d'autres éléments. N[sujet] +V+ N[COD] //dans/de N</p>	accompagner_2, combiner_2
	<p><b>cause_objectif</b> Relation causale avec indication du point d'aboutissement. N[sujet] +V+ à N[COI] N[sujet] +V+ sur N[COI]</p>	amener, inciter, renvoyer_1
	<p><b>correspondance</b> Relation caractérisant un état d'une entité par le rapprochement avec d'autres entités, mettant en évidence leurs ressemblances/conformité. N[sujet] +V+ N[COD]// à/avec N[COI]</p>	adapter_s'_2, ajuster, coïncider, converger_2, correspondre, satisfaire_2
	<p><b>implication_cause</b> Relation d'implication consistant en ce qu'une chose (un fait, une idée, un caractère) entraîne comme conséquence logique</p>	conditionner, déterminer_2, engendrer, entraîner_1, entraîner_2, évoquer_1, expliquer_2, générer_2, impliquer,



	<p>une autre. N[sujet] +V+ N[COD]/à N[COI]/que P</p>	<p>imposer_3, induire, motiver, produire_1, provoquer, sous- tendre, suggérer_2, supposer_3</p>
	<p><b>influence</b> Relation renvoyant au fait de produire un effet suscitant des modifications, manière d'agir sur qqn ou qqch. N[sujet] +V+ sur N[COI]</p>	<p>guider, influencer_2, inspirer_1, opérer_2</p>
	<p><b>opposition</b> Relation caractérisant un état d'une entité par le rapprochement avec d'autres entités, mettant en évidence leurs différences. N[sujet] +V+ N[COD]// de/avec /à N[COI]</p>	<p>affranchir_2, contraster, contredire, différencier_se_2, différer, distinguer_se_1, opposer_3, séparer_2</p>
	<p><b>source</b> Relation mettant en lien l'origine, la raison de qqch et le résultat. N[sujet] +V+ de N</p>	<p>découler, dépendre, hériter, procéder_2, provenir, ressortir_1, résulter</p>
<p><b>temporalité</b> Classe de verbes caractérisant un processus par rapport au temps : durée, fréquence, chronologie (localisation sur l'axe temporel), début et fin.</p>	<p><b>#début</b> N[sujet] +V+ N[COD]</p>	<p>amorcer</p>
	<p><b>#durée</b> N[sujet] +V+ N[COD]</p>	<p>entretenir_1, entretenir_2</p>
	<p><b>#fin</b> N[sujet] +V+ N[COD]/à N</p>	<p>aboutir_2, conclure_1</p>
	<p><b>chronologie</b> N[sujet] +V+ N[COD]/à N</p>	<p>précéder_1, succéder</p>
	<p><b>début</b> N[sujet] +V</p>	<p>apparaître_1, émerger, établir_s'_3</p>
	<p><b>durée</b> N[sujet] +V N[sujet] +V+ N[COD]</p>	<p>conserver, maintenir, subsister</p>
	<p><b>fin</b> N[sujet] +V+ à/vers N[COI]</p>	<p>aboutir_1, contribuer, converger_1, mener_1, résumer_se_2, tendre_2</p>

---

## ANNEXE 3

### Expérimentation didactique I

Cette expérimentation devrait permettre aux étudiants (avec un niveau au minimum B1) de s'initier au discours scientifique : l'accent sera surtout mis sur les opinions de l'auteur.

#### I. Séquences générales sur l'écrit scientifique et le LST

#### II. Séquences particulières sur les constructions verbales des verbes d'opinion

Dans l'écrit scientifique, il existe des formules presque figées que l'auteur utilise pour structurer le discours (*dans un premier temps, autrement dit, etc.*), annoncer ses hypothèses de recherche, sa problématique (*nous nous demandons si..., nous formulons l'hypothèse que..., etc.*), exprimer son opinion (*nous considérons que..., nous rejetons l'approche X..., etc.*). Dans les séances suivantes, nous allons travailler sur les verbes d'opinion (*Admettre, affirmer, estimer, juger, penser, croire, considérer, opposer, rejeter, contester, exclure, supposer*) et leurs constructions.

#### Séquences générales sur l'écrit scientifique et le LST

##### [1] Découverte de l'écrit scientifique

**Consigne : Lisez et comparez les deux documents ci-dessous et répondez aux questions.**

---

Document 1 :

Un nouveau titre, et pas n'importe lequel, sur un curriculum vitae déjà très fourni : 24 pages de distinctions, de publications, de prix en tous genres, auxquels s'ajoute donc, depuis lundi 13 octobre, ce prix Nobel d'économie 2014. Jean Tirole, chercheur à l'université de Toulouse et directeur de l'école d'économie de Toulouse, était invité de RFI mardi 14 octobre.

RFI : Se réveiller Nobel, cela change quoi au juste ?

Jean Tirole: J'espère que cela ne va pas changer trop de choses, je n'ai pas très bien dormi cette nuit mais c'est merveilleux ce qui m'arrive. Je pense que ça va être un peu difficile à vivre dans le court terme, mais j'espère pouvoir retrouver ma vie de chercheur. Ce n'est que du bonheur, en tout cas, pour l'instant. [...]

RFI: Un Nobel dans un pays en crise, cela fait dire à certains commentateurs qu'il y a peut-être trop de distance, chez nous, entre théorie et pratique. Partagez-vous ce point de vue, à savoir que les Français sauraient mieux parler économie que la gérer ?

Je pense qu'il faut augmenter et améliorer la culture économique en France. Certes, elle n'est développée nulle part dans le monde, mais particulièrement, en France, je crois qu'il y a des progrès à faire. Dans l'enseignement évidemment au lycée, dans les universités, dans les grandes écoles, pour essayer de mieux comprendre les enjeux, c'est vraiment un point important.

*Jean Tirole : « Plus de compromis politiques pour des réformes durables » par Caroline Paré, Francefinances, publié le 14/10/2014 et modifié par Yan le 15/10/2014*

1. De quel type de document s'agit-il ?  
un reportage B. un portrait C. un interview D. une critique
2. Quel est le point de vue de Jean Tirole par rapport à la culture économique ?  
Relevez les mots qui vous permettent de justifier votre opinion.

Document 2 :

Pour tenter d'expliquer pourquoi les femmes ingénieurs accèdent moins souvent que leurs collègues masculins aux fonctions dirigeantes, nous avons mené une recherche auprès d'élèves ingénieurs masculins et féminins poursuivant leur formation à l'Institut National des Télécommunications. Nous nous sommes demandé si les projets professionnels et personnels respectifs à chacun des deux sexes se ressemblent ou pas en première année. En d'autres termes, les étudiantes envisagent-elles de rentabiliser leur diplôme d'ingénieur de la même façon que les étudiants ? Les résultats de l'étude montrent que les projets d'avenir des étudiantes sont assez différents de ceux de leurs homologues masculins. Elles ont, en effet, globalement des projets professionnels moins ambitieux qu'eux car pour elles vie professionnelle et vie familiale entreront souvent en compétition. Elles envisagent notamment plus souvent que les garçons des aménagements de leur travail pour dégager du temps pour leur vie familiale. On peut penser qu'elles anticipent de « s'adapter raisonnablement », selon l'expression de M. Duru Bellat (1990), à la division sexuée du travail mais aussi qu'elles ont des difficultés à se projeter dans des rôles pour lesquels il existe encore pour l'instant bien peu de modèles féminins. On retrouve là le poids des rôles de sexe. Les femmes, quelle que soit leur formation, sont plus lourdement mises à contribution que les hommes pour assurer la marche de la structure familiale et payent le prix de ce rôle en matière de promotion et de carrière.

*Christine Fontanini « Elèves ingénieurs : aspirations et projets de vie. Une question de genre. », Scientext communication\_142\_sciedu\_AREF2007\_Fontanini\_entete*

1. De quel type de document s'agit-il ? (choix multiples)  
A. un document littéraire B. un document scientifique C. un document journalistique D. un document descriptif E. un document informatif F. un document argumentatif
- (1) Quel est l'objectif de l'étude menée par l'auteur ?
- (2) Quelles sont les questions posées par l'auteur ?
- (3) Quels sont les résultats obtenus de l'étude ?
2. Quel est le point de vue défendu par l'auteur ? Relevez les expressions où se trouve l'opinion. Connaissez-vous d'autres verbes ou d'autres expressions qui expriment l'opinion ? Cherchez dans un dictionnaire les synonymes des verbes (ou des expressions) puis refaites la phrase avec ceux de votre choix.
3. Comparez les expressions de l'opinion relevées avec celles du document 1, quelles sont les différences ? Essayez d'expliquer ces différences.
4. D'après vous, cet extrait vient de quelle partie de l'article ? Et pourquoi ?  
A. résumé B. introduction C. théorie D. méthodologie E. conclusion
5. Trouvez un synonyme de l'expression « mener une recherche ».

---

## [2] Observation de l'écrit scientifique

**Consigne : Lisez le document 3 et répondez aux questions**

Document 3 :

À partir de la description statistique des pratiques de mobilité des étudiants internationaux, la première partie de cet article met en évidence plusieurs profils de mobilité qui se cristallisent autour de la variable " continent d'origine ". Dans un second temps (Partie 2) et grâce à l'analyse des données contextuelles et des entretiens, nous montrons que cette variable n'explique rien en tant que telle, cachant en réalité des effets de structure de populations étudiantes très différentes qui ne profitent pas des mêmes conditions de séjour et qui ne détiennent pas les mêmes ressources sociales. [geo-art-6-introduction]

1. Qu'est-ce que l'auteur souhaite expliquer à travers cet extrait (sa fonction dans l'article), cet extrait vient de quelle partie de l'article (résumé, introduction, conclusion) ?
2. Dans l'article, qu'est-ce que l'auteur présente dans un premier temps ?
3. Dans la partie 2 de l'article, que pense l'auteur de la variable "continent d'origine" ?

## [3] Découvrir les différentes parties de l'article scientifique

**Consigne : Identifiez les différentes parties d'un article scientifique et associez chaque extrait à la partie à laquelle il appartient**

A. introduction B. théorie C. conclusion D. résumé E. méthodologie

1) A l'issue des analyses effectuées, quelques points méritent d'être soulignés à l'égard des représentations et des pratiques langagières des apprenants franco-allemands. Nous remarquons, en premier lieu, que les représentations fonctionnent comme des mécanismes qui aident les élèves à gérer le choix des langues au cours de l'interaction et à interpréter leur sens et leur valeur conversationnelle.

2) En partant de ces observations, nous allons tenter de comprendre, par une analyse empirique, ce qu'est une séquence médiatique « chargée d'émotions ».

3) Pour apporter des éléments de réponse à la question de la réception, nous avons mis en place un « test de réception », qui ne tient pas lieu de véritable étude de la réception mais qui donne quelques indications pertinentes pour notre analyse. Nous avons projeté la séquence décrite à trente étudiants (inscrits en formation d'ingénieurs à l'Université de Technologie de Troyes). Ces étudiants suivent un cours d'analyse des médias et ont répondu au test dans le cadre d'une séance de travaux dirigés.

---

4). On pourra donc parler de « dispositif émotionnel » pour définir cette combinaison de paramètres et faire l'hypothèse qu'une séquence de parole à la télévision sera d'autant plus chargée émotionnellement qu'elle fera intervenir un grand nombre d'éléments de ce dispositif.

5). Celle-ci utilise divers procédés rhétoriques favorisant l'inférence émotionnelle (Plantin, 1998) : l'organisation thématique de son intervention (le désespoir et la perte d'emploi répétée), l'explicitation des émotions (« c'était pour parler désespoir »), la personnification (le discours est en « je »), la mention de détails (« un BTS secrétariat trilingue »), l'utilisation de procédés stylistiques d'amplification (« complètement désespérée »), des constructions en crescendo (« désespérée... démoralisée... aucun sentiment... morte »), le recours à la métaphore (« point d'interrogation »), l'interjection (« Bon Dieu »), la conclusion alexithymique (J. Cosnier, 1994, p. 140)...

6). Cet article traite des émotions dans une émission de télévision, et plus précisément de la mise en place d'un « dispositif émotionnel » au cours d'un « talk-show ». Ce travail est basé sur l'analyse multimodale de l'émission « Demain les jeunes », diffusée sur France 2 en mars 94. L'analyse de cette séquence émotionnelle (une jeune fille parlant de son désespoir) montre que sa force émotionnelle dépend d'un dispositif global et multimodal. L'expression des émotions est basée sur le verbal, le vocal et le mimogestuel, mais est aussi induite par un cadrage préalable (l'animateur parle de l'angoisse de la jeunesse) et renforcée par le filmage (gros plans). Cette analyse est complétée par un test de réception. Ce travail s'inscrit dans le champ de la pragmatique des interactions médiatiques.

#### **[4] Définir l'écrit scientifique**

**Consigne : A travers les exercices ci-dessus, donnez une définition de l'écrit scientifique. Quels sont les éléments importants dans l'écrit scientifique ? Et quels éléments vous permettent de le distinguer des autres types d'écrits (écrit littéraire, journalistique, etc.) ?**

#### **Proposition de la définition de l'écrit scientifique :**

L'écrit scientifique est un type d'écrit argumentatif, un écrit d'opinion qui permet de faire avancer la connaissance au sein d'une communauté scientifique. Il vise à décrire les recherches scientifiques menées par l'auteur dans un domaine particulier et à diffuser les résultats d'une recherche intégrant une synthèse critique de l'état du terrain de recherche avec la précision des objets d'étude, des objectifs et des hypothèses ; un cadre théorique articulé ; une présentation détaillée des méthodologies choisies, des analyses des données et des résultats obtenus ; le soulignement de l'opinion ; les suggestions et les limites de la recherche.

#### **I. Séquences particulières sur les constructions verbales des verbes d'opinion**

Dans l'écrit scientifique, on remarque que certains verbes sont employés pour exprimer une opinion de l'auteur (*penser, croire, considérer, estimer, juger*, etc.). L'opinion de

---

l'auteur peut être favorable (*admettre, affirmer*), neutre (*penser, considérer*) ou défavorable (*rejeter, contester*). Certains verbes d'opinion sont souvent modalisés ou introduits dans une subordonnée de condition (*on peut considérer que, si l'on admet l'idée, il est possible de supposer que*), ce qui atténue la prise de position individuelle. De ce fait, l'opinion de l'auteur est moins subjective et plus partagée.

### [1] Découverte de l'opinion

**Consigne : Lisez les extraits suivants et répondez aux questions**

#### Extrait 1 :

En suivant cette piste relative à l'influence des moyens de communication sur les modes de pensée, il est possible de considérer que - de la même manière que lors de l'introduction de l'écriture et de l'imprimerie - l'essor actuel des Technologies de l'Information et de la Communication (TIC) en général, mais aussi en formation, a et aura des répercussions majeures sur l'évolution des connaissances et des modes de raisonnement. (*Sciences de l'éducation : De la raison graphique à la raison numérique ? Evolutions du rapport à l'écrit dans les espaces virtuels*)

1. Quelle est l'opinion de l'auteur par rapport à l'essor actuel des TIC ? Relevez les expressions où se trouve l'opinion.
2. Parmi les verbes suivants, lequel (ou lesquels) pourrait-on utiliser avant la complétive « que » comme synonyme du verbe « considérer » ? *Admettre, affirmer, penser, croire, juger, estimer*. Justifiez votre choix à l'aide d'un dictionnaire.
3. Essayez de remplacer « il est possible de considérer que » par une autre construction équivalente.

#### Extrait 2 :

En effet, un peu plus de la moitié (57,1%) des interviewés préfèrent la musique aux paroles turques (8 sur 14, dont la moitié n'a mentionné que la musique turque et le reste l'a mise en parallèle avec la musique grecque) ; trois sur quatorze (21,4%) aiment la musique turque autant que la musique grecque, tandis que deux sur quatorze (14,2%) déclarent préférer la musique grecque et un seul interviewé (7,1%) prétend aimer plus la musique grecque, mais nous nous permettons de contester son affirmation, vu l'ensemble de son discours particulièrement surveillé et parfois contradictoire. (*Linguistique : La place et les fonctions réservées à la langue minorée : le cas de baфра (grèce)*)

1. Relevez les expressions où se trouve l'opinion de l'auteur. Son opinion est-elle favorable ou défavorable ?
2. Qu'est-ce que l'auteur conteste ? Et pourquoi ?
3. Parmi les verbes suivants, lequel pourrait-on utiliser auprès du nom « affirmation » comme synonyme de « contester » ? *Rejeter, opposer, exclure*, à votre avis, pourquoi certains ne sont-ils pas possibles ?

4. Pourquoi l'auteur dit « nous nous permettons de », il y aura quelle différence si on l'enlève de l'extrait ?
5. À qui renvoie le pronom « nous » dans l'expression « nous nous permettons de » ?

**Extrait 3 :**

L'hypothèse retenue est celle d'une modification définitive de l'environnement au-delà d'un seuil critique de pollution. Le choix est donc essentiellement binaire : soit l'on dépasse cette limite, jugeant que cette norme implicite est trop coûteuse comparativement aux dommages catastrophiques qu'elle permet d'éviter, soit l'on s'y conforme, considérant, au contraire, que la destruction de l'environnement due à la catastrophe justifie des mesures de dépollution drastiques. (*Economie : Catastrophe climatique irréversible et politique de l'effet de serre*)

1. Quelles sont les deux opinions opposées dans l'extrait ? Relevez les expressions où se trouve l'opinion.
2. Parmi les verbes suivants, lequel pourrait-on utiliser pour remplacer « jugeant » et « considérant » ? *Admettre, affirmer, penser, croire, soutenir, supposer, estimer, trouver.*
3. À qui renvoie le pronom « on » dans l'extrait ? (l'auteur, l'auteur et le lecteur, l'auteur et la communauté scientifique limitée, tout le monde, le lecteur, les autres)

**[2] Comprendre les sens des verbes d'opinion**

**Consigne : Identifiez les constructions verbales portant le sens d'opinion et associez les constructions verbales aux sens correspondants (voir aussi le *Trésor de la langue Française informatisé*).**

<p>1. En ce qui concerne la dénomination des types de mobilité, nous considérons qu'il existe trois critères de distinction.</p> <p>2. Les théories peuvent en effet être évaluées en considérant les phénomènes linguistiques qu'elles permettent de décrire ou d'expliquer.</p>	<p>a. prendre en considération ou tenir compte de quelque chose</p> <p>b. donner son jugement sur un objet ou un phénomène</p>
<p>1. Nous estimons pour notre part que cette perspective caricature à l'excès une situation bien plus complexe.</p> <p>2. De récentes données expérimentales nous permettent d'estimer l'évolution de la force de la mémoire haptique chez le jeune bébé au cours de la première année.</p> <p>3. On estime par ailleurs à environ 225 000 personnes la migration pendulaire de navetteurs non recensés.</p>	<p>a. donner une évaluation sur quelque chose d'abstrait</p> <p>b. donner une opinion sur un objet ou un phénomène</p> <p>c. prendre la mesure du prix, de la durée ou de la quantité de quelque chose en donnant une référence statistique</p>
<p>1. On peut en effet juger que l'assertion négative conserve la pertinence de l'énoncé.</p>	<p>a. être d'avis que</p> <p>b. prendre la mesure de quelque chose d'abstrait</p>



<p>2. Nous proposons une étude quantifiée du modèle, afin de juger de l'optimalité du ratio capital public/PIB en France.</p>	
<p>1. Finalement, à supposer qu'il existe bien une différenciation des aptitudes avec l'âge...  2. Tout d'abord, contrairement à ce que nous supposons, ZEIRA fait l'hypothèse que le bien produit dans le premier secteur sert à la fois de bien de capital et de bien de consommation.  3. Ces variations de l'indice laissent supposer que des phénomènes de différents types influencent le niveau de tolérance du public français.  4. Le total atteindrait 18,3 milliards, avec un rythme de trois milliards par an, ce qui suppose le maintien d'une croissance régionale rapide.</p>	<p>a. penser quelque chose comme probable sans pouvoir affirmer de façon positive  b. donner le moyen de se former une idée sans affirmation  c. impliquer une condition nécessaire ou un fait scientifique  d. poser à titre d'hypothèse sans jugement</p>
<p>1. Nous ne contestons pas la pertinence du modèle SILEX dans l'étude sémantique des mots suffixés, mais par contre nous questionnons sa pertinence dans le domaine préfixal .  2. On peut admettre que la méthode possède la propriété de préserver des mésinterprétations .  3. On peut néanmoins penser qu'elle aura un effet indirect ou différé en suscitant sur le tard chez l'auteur littéraire.  4. En revanche, on rejette systématiquement la non-causalité pour les données britanniques.  5. On pourrait croire qu'elles constituent l'enjeu ultime de la prime scolarité.  6. À ce titre, Mellon s'oppose au principe de l'impôt sur les ventes (sales tax) qu'il juge politiquement désastreux et peu convaincant en termes budgétaires.  7. Comme nous l'avons affirmé ci-dessus, seule une étude sur les routines pourrait nous dire si le feeling, la méthode, ou les deux (habitus) correspondent à cette sensibilité reproductive.  8. En fait, on peut exclure l'idée que les bénéfices perçus proviennent de bénéfices réels autres que ceux qui ont été investigués jusqu'à présent.</p>	<p>a. après un jugement de réalité, accepter quelque chose par l'esprit  b. se former dans l'esprit une idée prise sur quelque chose  c. un jugement que la personne a énoncé comme vrai plus haut  d. ne pas mettre en discussion ou en doute la réalité ou la vérité d'un fait, la valeur d'une idée, d'une théorie  e. refuser totalement quelque chose une idée, ou un fait</p>

### [3] Travailler sur les constructions syntaxiques et les verbes synonymes

**Consigne : Parmi les verbes proposés entre parenthèses, choisissez le synonyme (ou les synonymes) qui peut (peuvent) remplacer le verbe d'opinion dans les constructions verbales.**

1. On pourrait **penser** (considérer, estimer, admettre) résoudre la difficulté en opposant d'un côté les marques linguistiques relevant de l'énonciation, et en **rejetant** (s'opposer à, exclure) tous les autres critères dans la masse informelle de la stylistique.

- 
2. Le choix est donc essentiellement binaire : soit l'on dépasse cette limite, **jugeant** que (estimer, admettre) cette norme implicite est trop coûteuse comparativement aux dommages catastrophiques qu'elle permet d'éviter, soit l'on s'y conforme, considérant, au contraire, que la destruction de l'environnement due à la catastrophe justifie des mesures de dépollution drastiques.
  3. Pour "faciliter les changements qu'exige le progrès économique", les auteurs **estiment** (considérer, admettre) nécessaire de compenser les déséquilibres du marché dans deux domaines : le soutien aux régions sous-développées et la réadaptation des travailleurs.
  4. Plusieurs indices laissent **penser** que (supposer, considérer) ces alliances opèrent aux dépens des régulations professionnelles habituelles.
  5. Comme l'**affirment** (penser, croire) deux auteurs : " Le judiciaire structure ou donne les orientations nécessaires à la politique publique " (Cannon et O'Brien, 1985, 253).
  6. Prises en charge par le conseil, ces mesures ne permettent pas de **supposer** (penser, admettre) que l'initiative, ni même le reste de l'organisation, procèdent de la même instance.
  7. Quant au problème plus spécifique du rapport entre les emplois discursifs d'un mot et ses autres emplois, nous ne **pensons** pas (juger, croire) qu'il y ait de rupture nette.
  8. Nos résultats permettent de **rejeter** (s'opposer à, exclure) l'hypothèse selon laquelle la faiblesse de leur maintien dans le temps pourrait expliquer ce paradoxe.
  9. Nous ne **contestons** pas (rejeter, s'opposer à) la pertinence du modèle SILEX dans l'étude sémantique des mots suffixés, mais par contre nous questionnons sa pertinence dans le domaine préfixal.
  10. Le langage mathématique est en effet **jugé** (supposer, penser) supérieur au langage naturel (Simon, 1957 , 90).

#### [4] Travailler sur l'antonymie

**Consigne: Remplacez le verbe d'opinion par les antonymes de votre choix.**

**Plusieurs choix sont possibles.**

1. Pourtant, même si l'on **admet** (s'opposer à, contester, rejeter, exclure) l'idée d'une certaine universalité du traitement orthographique, la variété des tendances n'est pas négligeable pour autant.
2. C'est pourquoi ils **s'opposent au** (admettre, affirmer, accepter) naturalisme en sciences sociales, lequel leur semble, à juste titre, constituer l'orthodoxie des sciences de la cognition.

#### [5] Renforcer-Fixer

**Consigne : Lisez le texte suivant et répondez aux questions**

---

Un autre point important, sur lequel nous allons insister, est que nous considérons que les caractéristiques de ces instruments ne présagent en rien de l'activité qui va en découler pour l'utilisateur. Nous rejoignons Perriault, pour qui il faut admettre les usages détournés comme des indicateurs d'appropriation effective. Avec Pierre Schaeffer, nous considérons que toute nouvelle utilisation d'une machine à communiquer constitue une hypothèse à vérifier (Schaeffer, 1970). En d'autres termes, nous rejetons le postulat d'utilité universelle des machines relevant de cette catégorie : leur utilité est toujours à démontrer.

1. Quelles sont les opinions de l'auteur ? Relevez les mots qui vous permettent de justifier votre choix.
2. L'auteur est-il du même avis que Perriault ? Pourquoi ?
3. Quel synonyme pourriez-vous envisager pour remplacer les verbes d'opinions dans le texte ?
4. Réécrivez le texte avec les antonymes des verbes d'opinions.

**Consigne : Réécrivez les phrases suivantes avec les verbes ou les constructions données**

EX : Marie Duru Bellat (1990) montre que pour comprendre les choix d'orientation et professionnels des filles, il est nécessaire de prendre en compte le fonctionnement actuel de la famille. (juger)

Réécriture : Pour comprendre les choix d'orientation et professionnels des filles, Marie Duru Bellat (1990) juge nécessaire de prendre en compte le fonctionnement actuel de la famille.

1. Ces résultats nous poussent à croire qu'on trouvera, chez des adolescents, davantage de moyens de compensation, phénomène qui serait lié à une diminution des erreurs. (amener /conduire à penser que)
2. De façon générale, nous pouvons affirmer que le sens est aussi soumis à la norme prescriptive, puisque les différents sens associés aux vocables sont décrits dans les dictionnaires et sont plutôt statiques. (possible de penser que)
3. Comme la source des problèmes mentionnés plus haut est la méconnaissance de propriétés lexicales, nous croyons plus juste de parler d'erreurs lexicogrammaticales (ou lexicosyntaxiques dans le cas de problèmes de régime). (estimer)
4. Arrivée à ce point de notre discussion, nous croyons pouvoir formuler encore une remarque concernant les études sur l'objet direct/transitivité. (penser)
5. À partir de ce point de vue, on peut considérer le cri comme un modèle à partir duquel une variété d'actes illocutoires expressifs pourront être produits. (être considéré comme)
6. L'intérêt du focus group réside dans le fait de recueillir les opinions des personnes formulées dans le cadre d'une discussion, en pensant que la dynamique et les

---

échanges qui se créent entre les personnes les aident à développer leur réflexion.  
(considérer que)

**Consigne : A l'aide de la liste des mots, remplissez les trous par les mots corrects, n'oubliez pas de conjuguer les verbes si besoin.**

Juger, admis, si, considérer, croire, en revanche, penser, amener,

1. En linguistique, il est \_\_\_\_\_ de considérer que la langue répond à des règles que l'on peut expliciter.
2. À l'instar de Baron et Bruillard (2001) nous ne \_\_\_\_\_ pas qu'une approche centrée sur un champ théorique unique se suffit à expliquer la diversité des phénomènes à prendre en compte.
3. Ainsi l'introduction des OP permet de \_\_\_\_\_ l'enseignement comme une situation d'activité instrumentée dans laquelle l'usage de l'ordinateur constitue un des moyens d'action de l'enseignant dans son travail.
4. Cette recherche m'a \_\_\_\_\_ à considérer les processus d'étiquetage comme des composantes de ce qu'on appelle de manière indifférenciée les dynamiques identitaires.
5. Essentiellement discursive, cette ostension de soi intervient lorsque les personnes veulent changer des pratiques sociales (contrôle des activités, comportements, manières de considérer), précisément liées à cet étiquetage et qu'elles \_\_\_\_\_ insatisfaisantes.
6. Bien que peu surprenante, cette constatation est tout de même préoccupante \_\_\_\_\_ l'on considère que les textes ont été rédigés à la maison et que le recours aux ouvrages de référence était permis.
7. D'un point de vue théorique, nous \_\_\_\_\_ avoir démontré qu'une délimitation précise de la notion d'erreur lexicale est possible et que la TST offre un cadre théorique favorable à cet objectif.
8. Concrètement, il s'agit par exemple de retirer des séquences \_\_\_\_\_ non indispensables si l'apprenant est pris par le temps (afin qu'il puisse se synchroniser avec le groupe).
9. Néanmoins, et malgré ses limites intrinsèques, une description iconique du sens ne devrait pas être totalement \_\_\_\_\_, car elle représente un point de départ, un moyen d'évocation très fort, une sorte de tremplin vers la compréhension des mots.
10. La distribution restant stable, on peut penser, \_\_\_\_\_, que cela s'explique par une stabilité de la fonction des concepts.
11. Dans cet ensemble, nous avons retenu uniquement ceux qui présentent au moins trois occurrences dans chacun des corpus, en \_\_\_\_\_ que pour pouvoir caractériser le fonctionnement en contexte de ces unités, on ne peut se contenter d'un hapax ni de deux occurrences.

---

## [6] Réutiliser

**Consigne : Lisez le texte suivant et soulignez les mots où se trouve l'opinion. En vous basant sur les différentes opinions, rédigez en quelques lignes votre propre point de vue sur la communication interpersonnelle dans l'enseignement et l'apprentissage. Vous pouvez utiliser les expressions suivantes :**

A l'instar de X, nous pensons que...

On peut considérer que

On affirme que

Il nous semble que

Nous estimons nécessaire/important de...

À travers la place du tuteur/formateur au niveau socio-affectif, nous avons déjà abordé l'importance de la communication. Cependant, cette communication doit s'élargir à l'ensemble du groupe. Albero (2003) insiste sur le fait que « la multiplicité des échanges qu'ils soient centrés sur les contenus, les processus ou procédures ou bien sur la dimension psycho-affective des apprentissages, fait partie intégrante de la formation ». Elle ajoute même que « l'interaction humaine est indispensable à l'auto-structuration des connaissances et à la construction de l'image de soi en tant qu'apprenant efficace ». D'Halluin et Biolluz (2006) considèrent aussi que les interactions interpersonnelles sont indispensables pour accentuer les activités de type comparatif et conceptuel qui manquent dans l'apprentissage individuel. Gaberan et Leprince (2006) tiennent le même discours et précisent que le tuteur, en position de mentor, doit favoriser ces échanges interpersonnels. Quant à Grosjean (2006), elle affirme que le tuteur joue « un rôle fondamental dans la construction du lien social » et celui-ci, en favorisant « le sentiment d'appartenance à une communauté éducative, est indispensable à tout processus d'apprentissage et constitue un élément essentiel contribuant à briser l'isolement de l'étudiant » (Glikman, 2002).

---

## ANNEXE 4

### Expérimentation didactique II

#### Fiche pédagogique de la première séance (120m):

##### Tour de parole: (15m)

Activités : Les apprenants s'expliquent leur sujet de recherche deux par deux, le deuxième doit expliquer au groupe le sujet de l'autre.

Explication : Cette étape vise à faire connaissance et à se familiariser avec le sujet de recherche de chacun.

##### Brainstorming : sensibiliser à l'écrit académique (mémoire) :

(15)

Activités : Les étudiants sont amenés à exprimer leurs idées sur le type d'écrit du mémoire avec des mots-clés. Après la mise en commun, les idées reformulées, classées seront présentées sous une forme synthétique comme indiqué ci-dessous dans l'encadré.

Explication : Cette étape permet de vérifier les connaissances des apprenants sur l'écrit universitaire et de faire le point. Ensuite, les idées rassemblées aident à la construction d'un savoir de base sur le type d'écrit du mémoire.

- |   |
|---|
| <ul style="list-style-type: none"><li>- caractéristiques énonciatives : marques de personnes avec des pronoms (<i>on, nous, etc.</i>), marques de contexte temporelles (<i>jusqu'à présent</i>), marques de modalisation avec des verbes modaux et le conditionnel (<i>pouvoir, etc.</i>), citations</li><li>- caractéristiques argumentatives : démontrer, justifier et interpréter par des reformulations, des comparaisons, ou des propos illustratifs</li></ul> |
|---|

---

## Découverte du LST et des constructions de verbes du LST (20

m)

Activités : Il s'agit de la première lecture/observation du texte scientifique où les étudiants sont amenés à travailler sur deux extraits (un résumé et une introduction) en répondant aux questions posées.

Explication : L'activité de compréhension écrite et de sensibilisation permet à l'apprenant de prendre conscience de l'importance des formules presque figées dans l'écrit universitaire que l'auteur utilise pour structurer le discours, formuler des hypothèses, exprimer son opinion, etc.

Consigne : Lisez les documents suivants et répondez aux questions.

### Document 1 :

Pour tenter d'expliquer pourquoi les femmes ingénieurs accèdent moins souvent que leurs collègues masculins aux fonctions dirigeantes, nous avons mené une recherche auprès d'élèves ingénieurs masculins et féminins poursuivant leur formation à l'Institut National des Télécommunications. Nous nous sommes demandé si les projets professionnels et personnels respectifs à chacun des deux sexes se ressemblent ou pas en première année. En d'autres termes, les étudiantes envisagent-elles de rentabiliser leur diplôme d'ingénieur de la même façon que les étudiants ? Les résultats de l'étude montrent que les projets d'avenir des étudiantes sont assez différents de ceux de leurs homologues masculins. Elles ont, en effet, globalement des projets professionnels moins ambitieux qu'eux car pour elles vie professionnelle et vie familiale entreront souvent en compétition. Elles envisagent notamment plus souvent que les garçons des aménagements de leur travail pour dégager du temps pour leur vie familiale. On peut penser qu'elles anticipent de « s'adapter raisonnablement », selon l'expression de M. Duru Bellat (1990), à la division sexuée du travail mais aussi qu'elles ont des difficultés à se projeter dans des rôles pour lesquels il existe encore pour l'instant bien peu de modèles féminins. On retrouve là le poids des rôles de sexe. Les femmes, quelle que soit leur formation, sont plus lourdement mises à contribution que les hommes pour assurer la marche de la structure familiale et payent le prix de ce rôle en matière de promotion et de carrière.

Fontanini Christine, « Élèves ingénieurs : aspirations et projets de vie. », Carrefours de l'éducation 1/2001 (n° 11) , p. 52-65

1. De quel type de document s'agit-il ? (choix multiples)  
A. un document littéraire B. un document scientifique C. un document journalistique D. un document descriptif E. un document informatif F. un document argumentatif
2. Découpez l'extrait en paragraphes et justifiez votre choix en précisant les motifs linguistiques et thématiques de votre découpage (Quel est l'objectif de l'étude menée par l'auteur ? Quelles sont les questions posées par l'auteur ? Quels sont les résultats obtenus de l'étude ?)

3. D'après vous, cet extrait vient de quelle partie de l'article ? Et pourquoi ?

A. résumé B. introduction C. théorie D. méthodologie E. conclusion

**Document 2 :**

Partant de ces différents arguments, nous avons mené une recherche auprès d'enfants bilingues français/anglais pour tester les hypothèses suivantes.

Public

Tout d'abord, les enfants bilingues ont un développement métaphonologique plus précoce que leurs pairs monolingues, c'est-à-dire un développement métaphonologique en dehors du contexte de l'apprentissage de la lecture. Ainsi, nous postulons que les enfants bilingues feront preuve de capacités métaphonologiques avant d'apprendre à lire. De plus, du fait de l'interaction entre les deux langues acquises, les enfants bilingues ne témoigneront pas du même type de préférence pour les unités subsyllabiques que les enfants monolingues. Nous supposons qu'ils ne feront pas preuve de la même sensibilité phonologique que leurs pairs monolingues.

Hypothèses

L'objectif est donc de déterminer la nature des capacités phonologiques des enfants bilingues, comparativement à des enfants monolingues. En d'autres termes, l'acquisition précoce d'une deuxième langue nous semble susceptible d'avoir une incidence sur le développement phonologique.

Objectif

(Laurent, A. (2008). Le développement métaphonologique chez les enfants bilingues et monolingues. In M. Loiseau et al. (Eds.), *Autour des langues et du langage : perspective pluridisciplinaire* (pp. 217-224). Grenoble : Presses Universitaires de Grenoble.)

1. L'extrait appartient à quelle discipline ? S'agit-il de la même discipline que le document 1 ?
2. Découpez l'extrait en paragraphes et justifiez votre choix en précisant les motifs linguistiques et thématiques de votre découpage.

## Analyse des phénomènes linguistiques dans l'introduction

(35)

Activités : Cette étape est divisée en trois parties dont chacune traite une fonction particulière dans l'introduction: 1) présenter l'objet théorique, 2) problématiser l'objet théorique, 3) poser les questions.



---

Dans chaque partie, l'apprenant doit travailler sur des extraits courts et repérer des constructions ayant trait aux fonctions mentionnées ci-dessus. Les constructions ainsi que les différents procédés utilisés seront étudiés.

Explication : Il s'agit d'une étape intermédiaire qui aide l'apprenant à faire une progression propice à la compréhension des phénomènes linguistiques nouveaux dans l'introduction.

## Présenter l'objet théorique

Consigne : À partir des extraits ci-dessous, il est possible de déterminer le contenu global de l'article. Quel est-il ? Quels sont les différents procédés utilisés pour présenter l'objet de l'article ?

Extrait 1.

Dans cet article, nous nous intéresserons à l'évolution de l'activité sur la plate-forme pendant les deux premières années d'usage par les étudiants (2005 et 2006). Nous montrerons quels facteurs ont permis l'émergence d'une double dynamique : une dynamique d'écriture et une dynamique communautaire, les deux se trouvant dans une relation d'interdépendance réciproque.

(Renáta Varga, 2007, Évolution des usages d'une plate-forme collaborative, Conditions d'émergence d'une double dynamique communautaire et d'écriture, *Congrès International ARE (Actualité de la Recherche en Éducation et en Formation)*, Aug 2007, Strasbourg)

Extrait 2.

S'appuyant sur un repérage préalable des changements de pratiques enseignantes (Boilevin & Brandt-Pomares & Ranucci, 2005), cette recherche se propose d'étudier la question de l'enseignement d'un savoir disciplinaire dans une situation mettant en œuvre des ordinateurs portables. Il s'agit alors d'expérimenter des dispositifs d'enseignement en sciences physiques et en technologie dans lesquels l'ordinateur est présent dès la conception de la situation et ainsi de tenir compte du fait que le mode de transmission appropriation changeant, les apprentissages et l'activité enseignante peuvent être modifiés.

(Jean Marie Boilevin, Pascale Brandt-Pomares, 2007, Intégration d'ordinateurs portables dans l'organisation de l'étude en physique et en technologie, *Congrès International ARE (Actualité de la Recherche en Éducation et en Formation)*, Aug 2007, Strasbourg, France. pp.1-12)

Extrait 3.

Les progrès technologiques, les contraintes économiques et les réflexions pédagogiques sont à l'origine de nombreux changements dans le monde de l'éducation et de la formation. D'ailleurs, on s'aperçoit que les dispositifs et les modalités de formation évoluent assez vite tout en se nourrissant les uns des autres. Au niveau de la formation d'adultes, Haeuw (2004) observait qu'« il semble qu'on est passé de l'individualisation (dans les années 80) à la FOAD et au e-learning (dans les années 90/2000) pour arriver maintenant à l'individualisation des parcours de FOAD ». On peut même ajouter le concept de FOADA (Formation Ouverte et A Distance, Accompagnée) qui insiste volontairement sur l'accompagnement technique et pédagogique des apprenants. Nous nous intéressons ici à l'expérimentation de la préparation du concours de secrétaire médical grâce à la FOADA.

(Catherine Renoult, 2007, De la pédagogie personnalisée à la FOADA il n'y a qu'un pas, *Congrès International ARE (Actualité de la Recherche en Éducation et en Formation)*, Aug 2007, Strasbourg, France. pp.1-12)

### Problématiser l'objet théorique (intérêt de l'étude)

Consigne : Dans les extraits suivants, observez comment l'auteur montre que l'objet traité présente de l'intérêt. Soulignez les éléments linguistiques qui exposent l'objet théorique/l'objectif théorique et ceux qui montrent l'intérêt de l'étude.

Extrait 4.

« Cet article propose une description systématique des emplois et valeurs des indéfinis *n'importe qui/n'importe quoi/n'importe quel N*. les indéfinis de cette série, comme la plupart des indéfinis autres que l'article indéfini, n'ont pas – à notre connaissance – fait l'objet de travaux un tant ou soit peu systématiques.

(D. Paillard, 1997, *Langue française* 116)

Extrait 5.

« la réflexion portera plus particulièrement sur *quelque chose* et d'autres indéfinis, dont l'intérêt théorique dans le cadre de la valence n'a jamais été sérieusement examiné.

(B. Defrancq & D. Willems, 1998, *Travaux de linguistique* 35)

Extrait 6.

Ces deux objectifs, premiers dans la recherche d'aide à l'écrit universitaire, ont rapidement révélé des interrogations dont les réponses restent encore en suspens malgré les efforts de chercheurs à ce propos : comment aborder ces collocations spécifiques auprès de publics universitaires non natifs du français (Verlinde et al., 2006) ? Quel type de description linguistique peut aider l'enseignant à mieux cerner les spécificités

de ces éléments linguistiques parfois complexes (au plan syntaxique et lexical notamment) (Leeman, 2005, Tutin, 2007a) ? Faut-il ériger ces collocations en « normes » scientifiques dès l'instant où elles apparaissent presque systématiquement (pour certaines) dans les écrits scientifiques (Binon et al., 2003) ?

Ces questions, vastes, guident notre réflexion afin de comprendre ce dont le didacticien a besoin au plan linguistique pour aborder sereinement ces éléments en cours de FLES.

(Cavalla Cristelle, Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE, 2015)

Extrait 7.

« Pour éviter une démarche aussi paradoxale, nous nous proposons, dans cet article de remettre en cause le postulat syntaxique dont elle procède et d'explorer une autre approche théorique de la composition nominale. A cet effet, nous nous efforcerons de montrer, à partir de l'examen d'un cas précis, <...> que la composition nominale met en œuvre des éléments et des modes d'assemblages spécifiques...(I. Tamba & B. Bosredon, 1991, Langue française 91)

Extrait 8 :

« Mon intention est de montrer que le nom propre n'est pas une sous-catégorie du nom, contrairement à ce qu'affirment un grand nombre de grammairiens, et d'essayer de comprendre pourquoi cette sous-catégorisation s'est si longtemps maintenu dans la tradition grammaticale. »

(N. Flaux, 1995, in M. Noailly (éd.), *Nom propre et nomination*, Paris, Klincksieck)

## Poser les questions

Consigne : Dans les extraits suivants, observez comment l'auteur pose des questions. Soulignez les éléments linguistiques qui indiquent les questions posées.

Extrait 8.

À partir de cette brève description de la situation linguistique de l'amazighe au Maroc, il semble légitime de s'interroger sur la réception de la langue nouvellement unifiée. Plus précisément, il est possible de formuler cette interrogation comme suit : l'unification des variétés dialectales amazighes peut-elle aboutir à la création d'une diglossie amazighe, avec une langue « haute », la langue standard, unifiée, qui serait réservée à l'écrit et aux domaines formels, et plusieurs variétés – les différents vernaculaires – qui demeureraient « basses » ?

(Myriam ABOUZAÏD, 2007, La standardisation de l'amazighe (berbère) au Maroc : enjeux et réception auprès d'enseignants, in M. Loiseau, M. Abouzaïd, L. Buson, C. Cavalla, A. Djaroun, C. Dugua, et al. (éd.), *Autour Des Langues Et Du Langage* :

*Perspective Pluridisciplinaire*. Colloque International Des Étudiants Chercheurs En Didactique Des Langues Et En Linguistique (p. 259-266). Grenoble: Presses Universitaires de Grenoble.)

Extrait 9.

A partir de ces éléments, nous nous sommes demandé si les étudiantes envisagent de rentabiliser leur diplôme d'ingénieur à plus ou moins long terme de la même façon que les étudiants. En d'autres termes, est-ce que les filles, en début de formation dans l'ingénierie, anticipent un moindre investissement de leur future vie professionnelle comparativement aux garçons? Pour tenter de répondre à cette question, nous avons mené une recherche auprès d'élèves ingénieurs masculins et féminins de l'Institut National des Télécommunications (INT) pour observer si les projets professionnels et personnels respectifs à chacun des deux sexes différent en première année de formation.

(Fontanini Christine, « Élèves ingénieurs : aspirations et projets de vie. », Carrefours de l'éducation 1/2001 (n° 11) , p. 52-65Extrait 10.)

## Exercice de remise en ordre (15m)

Activités : L'apprenant doit remettre en ordre des paragraphes d'une introduction en précisant les fonctions associées.

Explication : Après des activités de découverte des propriétés syntaxiques et sémantiques des constructions, l'apprenant entre dans la phase d'appropriation des éléments linguistiques. Cette activité leur permet de systématiser les fonctions des constructions étudiées.

**Consigne :** Les paragraphes de l'introduction suivante sont dans le désordre. Remettez-les en ordre en vous aidant des marques typiques de l'introduction et indiquez les différentes fonctions de cette introduction.

1.

L'analyse s'appuie sur des hypothèses psychologiques relatives à l'organisation des connaissances en mémoire (Rosch, 1978). Les connaissances sont structurées en catégories cognitives à partir de l'expérience sensible individuelle (Dubois, 2000). Ces catégories relèvent de la perception sensorielle dans différentes modalités sensibles. Elles peuvent être caractérisées par des propriétés descriptives. Le partage de propriétés entre plusieurs concepts permet d'identifier des « ressemblances de famille » et des représentations plus ou moins typiques au sein de catégories aux frontières floues.

2.

La première question que l'on se pose alors est la suivante : le confort en train est-il une catégorie cognitive de ce type ?

3.

L'analyse linguistique que nous proposons s'inscrit en réponse à une demande de la SNCF relative à la perception du confort global en train. Les objectifs du projet sont d'identifier les propriétés sémantiques du confort en train et leurs relations de dépendance, à partir d'analyses linguistiques et cognitives.

4.

Cet article porte sur la manière dont les formes linguistiques en contexte, utilisant les ressources de la langue mises en discours, renseignent sur les structures cognitives construites à partir des perceptions sensorielles. On tente ici d'identifier, à partir de l'analyse des formes adjectivales, les représentations individuelles et partagées qui se construisent dans les discours des voyageurs, lorsqu'on les interroge sur leur expérience sensible du confort.

(Gaëlle Delepaut, 2007. Contribution de la linguistique cognitive à l'identification du confort : analyse des discours des passagers sur le confort en train. Linguistics. Université de la Sorbonne nouvelle - Paris III.)

## Classement sémantique (10m)

Consigne : Classez les formules suivantes dans le tableau par groupe de deux. Les premiers exemples sont indiqués.

Activité : L'apprenant doit classer les différentes constructions verbales étudiées selon leurs valeurs sémantiques.

Explication : Cette étape vise à fixer les acquis de l'apprenant et leur permet de faire un récapitulatif. De plus, il peut approfondir ses connaissances sur les différents sens des verbes étudiés.

- 1) Dans cet article, on s'efforce (ra) de...
- 2) Dans cette perspective, cet article propose de...
- 3) On s'intéresse à la question...
- 4) Les questions que nous nous posons sont les suivantes...
- 5) On tente de...
- 6) Nous formulons/posons les questions suivantes...
- 7) Nous nous focalisons sur la question de...
- 8) Nous nous proposons d'étudier...
- 9) Nous nous sommes demandé si...
- 10) Nous cherchons à étudier...
- 11) Dans cet article, nous focaliserons notre propos sur la problématique de...
- 12) Cette étude vise à...

- 13) Nous abordons la question....
- 14) Cet article porte sur...
- 15) On s'attache à étudier...
- 16) Nous portons notre attention sur...

Présenter le sujet de recherche	Présenter l'objectif	Poser des questions
On s'intéresse à la question...	Dans cet article, on s'efforce (ra) de...	Les questions que nous nous posons sont les suivantes...

## VII. Production écrite (20m)

Activités : L'apprenant doit rédiger après le cours une quinzaine de lignes sur son sujet de recherche en utilisant les constructions étudiées. Un aide-mémoire qui récapitule toutes les constructions étudiées sera proposé.

Explication : Cette étape vise à consolider les connaissances acquises de l'apprenant.

## Fiche pédagogique de la deuxième séance (120m):

### I. Découverte du lexique et des constructions qui portent sur le positionnement (20m)

Activité : Dans cette étape, les étudiants seront amenés à découvrir le lexique et les constructions qui portent sur le positionnement. À partir d'un document écrit, ils doivent relever les deux points de vue de l'auteur et relever les arguments qui contribuent à les soutenir en soulignant les éléments linguistiques du positionnement.

Explication : Il s'agit de faire comprendre la notion du positionnement dans l'écrit scientifique et d'amener l'apprenant à découvrir les verbes du positionnement qu'on souhaite aborder dans un deuxième temps.

Consigne : Lisez le document 1 et répondez aux questions.

#### *Document 1*

Nous considérons premièrement l'apprentissage comme un processus situé. En ce sens, notre approche peut être catégorisée comme une approche praxéologique de l'apprentissage. En effet, nous considérons que l'apprentissage professionnel peut être étudié comme un processus qui prend place dans et à travers des actions situées : il est lié à des circonstances matérielles et sociales, elles-mêmes constituées historiquement et culturellement. Le processus d'apprentissage est ainsi lié à l'activité dans laquelle il prend place et se développe. De plus, le processus d'apprentissage émerge des interactions entre les acteurs et leur environnement social.

Nous empruntons ces idées, qui ont été discutées par de nombreux auteurs, aux champs de la *situated cognition* et du *cognitive apprenticeship* (Brown *et al.*, 1989 ; Lave et Wenger, 1991 ; Wenger, 1998) qui considèrent que la cognition et l'apprentissage sont situés et que l'activité est une part intégrante des processus d'apprentissage. Lave et Wenger (1991) soulignent à ce propos que l'apprentissage résulte de la participation à l'activité d'une communauté de pratique. Les activités situées s'avèrent être ainsi un analyseur fécond des processus d'apprentissage. <...>

Nous considérons aussi l'apprentissage comme un processus collectif. Nous nous fondons en cela sur les acquis de la psychologie du développement, dans la lignée de Vygotsky, qui admet que les capacités cognitives dites supérieures, ici l'acquisition de savoirs et de savoir-faire, ne se conçoivent pas strictement comme des mécanismes psychiques individuels, mais impliquent des processus sociaux, historiquement et culturellement ancrés.

Dans le contexte de la formation professionnelle qui nous intéresse ici, nous considérons que si la construction de connaissances et le développement de compétences sont individuels, les conditions et les ressources pour l'apprentissage sont sociales et collectives. Nous nous référons ici aussi aux travaux de Lave et Wenger (1991), pour lesquels l'apprentissage n'est pas une activité individuelle, mais résulte de la participation ratifiée à des communautés de pratique. Selon eux, une communauté de pratique peut être définie comme un ensemble de relations entre des personnes, des activités et le monde. Elle constitue une condition intrinsèque de la construction des savoirs. La participation à une communauté devient ainsi un principe épistémologique d'apprentissage. Dans le champ de la *situated cognition*, Rogoff (1996) montre également que l'apprentissage est lié à une communauté et constitue un processus de transformation de la participation à des activités socioculturelles.

Dans la conception de l'apprentissage comme participation à une communauté de pratique, une place importante est accordée à l'autre. Billet (2001), dans le champ du *workplace learning*, et Mayen (2002), dans le champ de la didactique professionnelle, soulignent le rôle central joué par l'autre dans les processus d'apprentissage. Ils insistent sur la nécessité d'étudier les ressources collectives accessibles aux travailleurs, qu'ils soient novices ou expérimentés, sur leur lieu de travail. Billet, en particulier, montre à quel point les membres plus experts d'une communauté professionnelle jouent un rôle de « guidage » (guidance) dans la facilitation de l'accès du novice à ces ressources.

(Barbara Duc. (2007). Apprentissage, participation à l'activité et identité professionnelle. Congrès international AREF 2007, actualité de la recherche en éducation et en formation.)

1. Relevez les deux idées principales soutenues par l'auteur et remplissez la colonne gauche du tableau suivant. Soulignez les éléments linguistiques qui vous permettent de justifier votre choix.
2. Relevez les chercheurs ainsi que leurs opinions qui servent d'arguments pour soutenir les deux idées de l'auteur et remplissez la colonne droite du tableau suivant. (le premier exemple est donné).

Idées principales de l'auteur	Chercheurs/opinions servant des arguments
1.	Lave et Wenger (1991) soulignent à ce propos que l'apprentissage résulte de la participation à l'activité d'une communauté de pratique.
2.	

3. Trouvez des expressions synonymes pour dire « nous considérons que... » à l'aide de l'outil CNTRL en ligne : <http://www.cnrtl.fr/synonymie/>

## II. Analyse des phénomènes linguistiques liés au positionnement (40m)

Activité : Cette étape est divisée en trois champs sémantiques : prise de position neutre/favorable/défavorable. L'apprenant sera amené à travailler sur des extraits courts et repérer des constructions qui expriment les différentes prises de position de l'auteur/du chercheur.

Explication : Il s'agit d'une étape intermédiaire qui aide l'apprenant à faire une progression propice à la compréhension des phénomènes linguistiques nouveaux par rapport à la prise de position, en particulier dans l'état de l'art.

Consigne : relevez les expressions qui portent sur la prise de position. Est-ce que l'auteur/le chercheur est d'accord avec ?

### Présenter une prise de position neutre

Extrait 1.

Nous considérons, à ce titre, que, dans la construction de la compétence plurilingue chez les élèves, la représentation d'une proximité affective entre les langues est plutôt privilégiée au détriment d'une représentation objective de la distance typologique entre les langues. Il est difficile de déterminer, à ce stade, si les représentations influencent le processus d'acquisition des langues chez l'apprenant. En revanche, nous pouvons affirmer que les représentations contribuent à la construction progressive d'une conscience métalinguistique sur les liens entre les langues, comme nous le constatons par le biais des extraits analysés au long de cet article. Enfin, ces quelques considérations nous amènent à dire que l'emploi de plusieurs langues dans l'interaction n'est pas un phénomène aléatoire, mais qu'il assure au contraire des fonctions pragmatiques dans l'élaboration du discours.

([lin-com-33-body] – Sophia, Stratilaki. (2008). Compétence plurilingue, représentations des langues et construction des connaissances chez les apprenants franco-allemands. *Autour des langues et du langage: perspective pluridisciplinaire*. pp.331-338.)



---

Extrait 2.

Le choix est donc essentiellement binaire : soit l'on dépasse cette limite, jugeant que cette norme implicite est trop coûteuse comparativement aux dommages catastrophiques qu'elle permet d'éviter, soit l'on s'y conforme, considérant, au contraire, que la destruction de l'environnement due à la catastrophe justifie des mesures de dépollution drastiques.

([eco-art-648-summary] – Sébastien Rouillon. (2000). Catastrophe climatique irréversible et politique de l'effet de serre. *Annales d'économie et de statistique*, N° 59.)

Extrait 3.

La libre circulation des travailleurs et des capitaux serait requise afin de tirer pleinement parti du potentiel du marché commun. Pour « faciliter les changements qu'exige le progrès économique », les auteurs estiment nécessaire de compenser les déséquilibres du marché dans deux domaines : le soutien aux régions sous-développées et la réadaptation des travailleurs.

(Denord François, Schwartz Antoine, « L'économie (très) politique du traité de Rome », *Politix* 1/2010 (n° 89) , p. 35-56)

Extrait 4.

On admettra donc qu'en matière de rémunération, l'accès à la profession enseignante n'a constitué qu'une promotion relative pour la majorité de ces formateurs dans la mesure où ils sont plus souvent issus des classes moyennes que des classes populaires. De ce point de vue la différence avec leurs parents se joue peut-être autant sur la sécurité de l'emploi que sur la rémunération. En revanche, on peut supposer que l'accès à la profession enseignante constitue pour des acteurs issus majoritairement des classes sociales à capital culturel modeste une promotion symbolique significative en raison du prestige intellectuel associé au statut d'enseignant...

([sed-art-412-body] - Des identités professionnelles à l'épreuve du changement : les formateurs de l'enseignement technique et professionnel en IUFM (Troger) (*Revue Française de Pédagogie*))

**Présenter une prise de position favorable par rapport à d'autres théories existantes**

Extrait 5.

Un autre point important, sur lequel nous allons insister, est que nous considérons que les caractéristiques de ces instruments ne présagent en rien de l'activité qui va en découler pour l'utilisateur. Nous rejoignons Perriault (op. cit., p. 6), pour qui il faut « admettre que les usages de ces machines ne se conforment pas toujours à une utilisation canonique de référence, en d'autres termes, admettre et retenir les usages détournés comme des indicateurs d'appropriation effective.

(François-Xavier Bernard. L'impact cognitif des dispositifs médiatiques sur les enfants d'âge préscolaire en situation d'apprentissage avec un adulte. Étude d'un cas de simulateur informatique dans le contexte d'une exposition scientifique. Education. Université René Descartes -Paris V, 2006.)

#### Extrait 6.

M. Merleau Ponty (1945), s'accorde avec Husserl et Heidegger pour dire que la temporalité en tant que « flux temporel », n'est pas un objet de notre savoir, mais une dimension de notre être. La relation du sujet incarné au monde n'est pas de l'ordre de la mise à distance réflexive mais relève de l'immersion.

(Marie Carcassonne. (2004). De la temporalité vécue à la temporalité racontée : hypothèses et observations concernant le développement de la mise en mots du temps chez l'enfant)

#### Extrait 7.

En effet, on se rapproche aujourd'hui plus de la conception générale des buts proposée par Kruglanski (1996), considérant les buts d'accomplissement comme un ensemble de variables (en relation à l'accomplissement), reliées entre elles de manière à former une sorte de système organisé qui donnerait une orientation générale face à la tâche et la situation académique.

([psy-the-504-body] - Conflit sociocognitif et buts d'accomplissement : effets interactifs sur l'apprentissage et le mode de régulation du conflit (Darnon))

#### Extrait 8.

En effet, à l'instar de C. Perelman (2002), nous pensons que ce sont les moments discursifs et délibératifs qui confèrent une rationalité aux activités pratiques. Nous convenons cependant qu'une évaluation plus fine mériterait d'être entreprise pour corroborer ou non cette conviction. Après tout, ce qui est dit ne l'est que pour quelques élèves ; qu'en est-il des autres ?

([sed-com-130-body] – Janner-Raimondi, Martine. (2007). Attitudes juridiques dans les conseils d'élèves du primaire. *Congrès International ARE (Actualité de la Recherche en Éducation et en Formation)*, Aug 2007, Strasbourg)

---

## Présenter une prise de position défavorable avec une mise en discussion

Extrait 9.

Les deux méthodes s'accordent à **rejeter l'hypothèse** alternative d'endogénéité, si l'on spécifie correctement le modèle. Cette concordance est d'autant plus remarquable que les tests fondés sur les régressions auxiliaires sont fréquemment critiqués pour leurs mauvaises propriétés sur petit échantillon (DAVIDSON, MAC KINNON [1990]).

Stéfan Lollivier. (2001) Endogénéité d'une variable explicative dichotomique dans le cadre d'un modèle probit bivarié. Une application au lien entre fécondité et activité féminine. *Annales d'économie et de statistique*, N° 62.

Extrait 10.

Nos analyses typologiques nous ont conduits à postuler un certain nombre de tendances structurelles et organisationnelles qui confirment, prolongent ou **remettent en cause des résultats** d'études antérieures sur la syllabe (chapitre III).

Isabelle Rousset. (2004). Structures syllabiques et lexicales des langues du monde. Données, typologies, tendances universelles et contraintes substantielles. Linguistics. Université Stendhal – Grenoble III.

Extrait 11.

**L'AC considère ainsi que** l'étude comparée des systèmes phonologiques, syntaxiques, morphologiques etc. de deux langues permet de prévoir les erreurs potentielles d'un individu natif d'une langue 1 apprenant une langue 2. La faiblesse du postulat de ce domaine a rapidement été soulignée puisque peu d'erreurs supposées apparaissent effectivement dans les productions. **En s'opposant au postulat fondateur de l'AC**, l'AE s'est intéressée aux erreurs produites.

GONAC'H, J. (2008), " Les déterminants dans l'interlangue des bilingues turc-français et turc-anglais ", in M. Loiseau et al. (Coord.), *Autour des langues et du langage*, Grenoble : Presses Universitaires de Grenoble, pp.291-198.

### III. Reformulation des phrases avec des mots donnés (10m)

Activité : Dans cette étape, l'apprenant doit réécrire les phrases avec des mots/expressions donnés. L'objectif est de faire manipuler les différentes constructions en matière de prise de position.

---

Explication : L'apprenant entre dans la phase d'appropriation des éléments linguistiques. Cette activité lui permet de systématiser les constructions étudiées.

Consigne : Réécrivez les phrases suivantes avec les verbes ou les constructions donnés.

EX : Marie Duru Bellat (1990) montre que pour comprendre les choix d'orientation et professionnels des filles, il est nécessaire de prendre en compte le fonctionnement actuel de la famille. (juger)

Reformulation : Pour comprendre les choix d'orientation et professionnels des filles, Marie Duru Bellat (1990) juge nécessaire de prendre en compte le fonctionnement actuel de la famille.

1. Ces résultats nous poussent à croire qu'on trouvera, chez des adolescents, davantage de moyens de compensation, phénomène qui serait lié à une diminution des erreurs. (amener /conduire à penser que)
2. De façon générale, nous pouvons affirmer que le sens est aussi soumis à la norme prescriptive, puisque les différents sens associés aux vocables sont décrits dans les dictionnaires et sont plutôt statiques. (possible de penser que)
3. Comme la source des problèmes mentionnés plus haut est la méconnaissance de propriétés lexicales, nous croyons plus juste de parler d'erreurs lexicogrammaticales (ou lexicosyntaxiques dans le cas de problèmes de régime). (estimer)
4. Arrivée à ce point de notre discussion, nous croyons pouvoir formuler encore une remarque concernant les études sur l'objet direct/transitivité. (penser)
5. À partir de ce point de vue, on peut considérer le cri comme un modèle à partir duquel une variété d'actes illocutoires expressifs pourront être produits. (être considéré comme)
6. L'intérêt du focus group réside dans le fait de recueillir les opinions des personnes formulées dans le cadre d'une discussion, en pensant que la dynamique et les échanges qui se créent entre les personnes les aident à développer leur réflexion. (considérer que)
7. Comme Baron et Bruillard (2001), nous ne considérons pas qu'une approche centrée sur un champ théorique unique se suffit à expliquer la diversité des phénomènes à prendre en compte. (à l'instar de, penser)
8. Pourtant, même si l'on rejette l'idée d'une certaine universalité du traitement orthographique, la variété des tendances n'est pas négligeable pour autant. (s'opposer à)

#### **IV. Production écrite à partir d'une liste des arguments**

**(15m)**

Activité : Dans cette étape, en petits groupes de 2-3, l'apprenant doit rédiger la suite d'un paragraphe à partir des arguments donnés. Il doit enchaîner les arguments des différents chercheurs avec les constructions étudiées.

---

Explication : Il s'agit toujours de la phrase d'appropriation des éléments linguistiques. Mais en même temps, cela prépare l'apprenant à rédiger un bout d'écrit sur la prise de position dans son propre mémoire.

Consigne : Écrivez la suite du paragraphe suivant à l'aide des arguments donnés en utilisant les verbes et les constructions de positionnement étudiées.

À travers la place du tuteur/formateur au niveau socio-affectif, nous avons déjà abordé l'importance de la communication. Cependant, cette communication doit s'élargir à l'ensemble du groupe...

---

---

---

---

Arguments :

Albero (2003): « la multiplicité des échanges qu'ils soient centrés sur les contenus, les processus ou procédures ou bien sur la dimension psycho-affective des apprentissages, fait partie intégrante de la formation ». Et puis, « l'interaction humaine est indispensable à l'autostructuration des connaissances et à la construction de l'image de soi en tant qu'apprenant efficace ».

D'Halluin et Biolluz (2006): les interactions interpersonnelles sont indispensables pour accentuer les activités de type comparatif et conceptuel qui manquent dans l'apprentissage individuel.

Gaberan et Leprince (2006) : le tuteur, en position de mentor, doit favoriser ces échanges interpersonnels.

Grosjean (2006) : le tuteur joue « un rôle fondamental dans la construction du lien social » et celui-ci, en favorisant « le sentiment d'appartenance à une communauté éducative, est indispensable à tout processus d'apprentissage et constitue un élément essentiel contribuant à briser l'isolement de l'étudiant ».

Proposition de réponse:

Albero (2003) insiste sur le fait que « la multiplicité des échanges qu'ils soient centrés sur les contenus, les processus ou procédures ou bien sur la dimension psycho-affective des apprentissages, fait partie intégrante de la formation ». Elle ajoute même que « l'interaction humaine est indispensable à l'autostructuration des connaissances et à la construction de l'image de soi en tant qu'apprenant efficace ». D'Halluin et Biolluz (2006) considèrent aussi que les interactions interpersonnelles sont indispensables pour accentuer les activités de type comparatif et conceptuel qui manquent dans l'apprentissage individuel. Gaberan et Leprince (2006) tiennent le même discours et précisent que le tuteur, en position de mentor, doit favoriser ces échanges

interpersonnels. Quant à Grosjean (2006), elle affirme que le tuteur joue « un rôle fondamental dans la construction du lien social » et celui-ci, en favorisant « le sentiment d'appartenance à une communauté éducative, est indispensable à tout processus d'apprentissage et constitue un élément essentiel contribuant à briser l'isolement de l'étudiant » (Glikman, 2002).

(Catherine Renoult « De la pédagogie personnalisée à la FOADA, il n'y a qu'un pas... » Congrès international AREF 2007)

## V. Classement des constructions selon l'intensité de la prise de position (10m)

Activité : L'apprenant doit classer les différentes constructions étudiées selon l'intensité de la prise de position.

Explication : Cette étape vise à fixer les acquis de l'apprenant et leur permet de faire un récapitulatif. En outre, l'apprenant peut approfondir ses connaissances sur les différents sens des verbes de positionnement qui s'organisent sur une échelle d'intensité.

Consigne : Classez les expressions suivantes dans le tableau selon le degré d'intensité dans la prise de position (de l'accord ++, vers le désaccord --).

1. On pense/considère/admet/affirme/estime/juge/souligne/suppose que...
2. On préconise/soutient/reconnaît que...
3. En considérant/pensant que...
4. Nous insistons sur le point/le fait...
5. On s'accorde avec qn...
6. On rejoint qn/ le point de vue de qn...
7. On se rapproche de (approche/théorie)...
8. On rejette l'hypothèse/la théorie...
9. On s'oppose au postulat...
10. On remet en cause les résultats/la théorie/l'idée...
11. On remet en question les résultats/la théorie/l'idée...
12. On conteste une hypothèse...

++	+	-	--
On s'accorde avec qn	On pense que	On remet en question	On rejette l'hypothèse

---

## VI. Production écrite (25m)

Activité : Après la séance, l'apprenant doit lire trois auteurs/ouvrages importants dans sa bibliographie. Il doit exprimer en une quinzaine de lignes sa prise de position par rapport à ces auteurs/ouvrages.

Explication : Le premier objectif est surtout de vérifier si les apprenants ont bien compris ce qu'ils ont appris en termes linguistiques. Ensuite, l'enseignant peut vérifier s'ils sont capables de comprendre les théories de la discipline.

Consigne : Dans le cadre de votre mémoire, vous allez choisir trois ou quatre auteurs/ouvrages importants ayant traité d'un même sujet de recherche. Rédigez en une quinzaine de lignes votre prise de position par rapport aux différents points de vue existants.

Pour commencer, on vous propose plusieurs expressions :

*En suivant la théorie/le principe de X..., on pense que*

*À l'instar de X, on considère que*

*On rejoint X/ le point de vue de X...*

*On se rapproche de (approche/théorie)...*

## Fiche pédagogique de la troisième séance (120m) : le choix de l'approche et la collecte des données

### I. Découverte du lexique et des constructions qui portent sur

#### le choix de l'approche et la collecte des données (20m)

Activité : Dans cette étape, à partir d'un document écrit, les apprenants travaillent en binômes pour repérer le lexique et les constructions qui portent sur le choix de l'approche et la collecte des données en répondant aux questions.

Explication : Il s'agit d'une étape de découverte qui permet à l'apprenant de comprendre et de se familiariser avec la façon dont on présente la méthodologie (choix, collecte des données).

Consigne : Lisez le document suivant et répondez aux questions.

Ayant pour objectif la mise au jour des préoccupations et des enjeux éthiques de la profession enseignante au collégial, tels qu'identifiés par les enseignants, nous avons opté pour la méthode du focus group. La méthode du focus group ou groupe de discussion (Geoffrion, 1992 ; Krueger, 1998 ; Morgan, 1997) permet de réunir des personnes dans un climat de discussion libre, par ailleurs animée et thématiquement

Explication du choix

orientée par un animateur/chercheur. Dans ce contexte, la discussion est ouverte et permet la formulation ou l'énoncé d'un large spectre d'enjeux, à partir de quatre ou cinq thèmes déterminés par le cadre théorique, selon un guide d'animation préétabli. Cette formulation ne se fait cependant pas en vase clos, mais dans le contexte d'une discussion qui favorise l'échange et la réflexion avec des pairs que ne permettent pas, par exemple, des entrevues réalisées sur une base individuelle.

Les données ont été recueillies auprès de six groupes de discussion dans lesquels les enseignants ont été réunis en fonction du secteur où ils enseignent, pré-universitaire, technique ou de la formation générale, à raison de deux groupes de discussion par secteur, dans des collèges différents, dont quatre étaient situés dans les grands centres urbains et deux en région.

([sed-com-444-body] - La mise au jour des enjeux éthiques de la profession enseignant (Gohier, Jutras, Desautels))

1. Quel est l'objectif de l'étude ?
2. L'auteur a utilisé quelle méthode pour atteindre cet objectif ? Et pourquoi le choix de cette méthode ? Soulignez l'élément linguistique qui justifie votre choix.
3. Quelle est la méthode de recueil des données ? Soulignez l'élément linguistique qui justifie votre choix.
4. Trouvez des expressions synonymes pour dire « opter pour » à l'aide de l'outil CNTRL en ligne : <http://www.cnrtl.fr/synonymie/>

## II. Analyse des phénomènes linguistiques liés à la méthodologie (40m)

Activité : Cette étape est divisée en trois champs sémantiques : le choix de l'approche, la limite d'un choix et la collecte des données. Les apprenants travaillent en groupe et seront conduits à explorer des extraits courts et repérer des constructions liées à ces trois champs sémantiques.

Explication : Il s'agit d'une étape intermédiaire qui aide l'apprenant à faire une progression propice à la compréhension des phénomènes linguistiques nouveaux par rapport au choix et à la collecte des données dans la méthodologie.

Consigne : Relevez les expressions qui portent sur le choix de la méthodologie. Relevez les expressions qui portent sur la limite d'un choix. Relevez les expressions qui portent sur la collecte des données.

### Présenter un choix favorable

Extrait 1.



Ce souci d'intégration des niveaux linguistiques, à travers l'objet que constitue la liaison, nous a conduite à adopter une démarche pluridisciplinaire, essentielle, nous semble-t-il, pour circonscrire notre champ d'étude. Cette pluridisciplinarité correspond à la fois à une ouverture à différents champs des Sciences du Langage (phonologie, syntaxe, etc.) et aux rapprochements que nous avons effectués entre linguistique et psychologie. Cette pluridisciplinarité nous est rapidement apparue nécessaire pour deux raisons.

[lin-the-505-introduction] - Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage (Dugua)

Extrait 2.

Dans un premier temps, nous avons transcrit les enregistrements des échanges à l'aide d'un tableau qui distingue les interventions du parent de celles de son enfant. Nous avons retenu un mode de transcription qui permet de rendre compte aussi bien des interventions verbales que des interventions non verbales (gestes, manipulations, mimiques, etc.) des deux partenaires.

[sed-the-69-body] - L'impact cognitif des dispositifs médiatiques sur les enfants d'âge préscolaire en situation d'apprentissage avec un adulte (BERNARD)

Extrait 3.

Pour mener à bien cet objectif, nous nous sommes appuyé sur un dispositif unique, l'élément " dessine avec ton doigt ", présenté dans l'exposition des 3/5 ans de la *Cité des enfants*, à la *Cité des sciences et de l'industrie*. Par ailleurs, nous avons privilégié une approche qualitative basée sur une démarche d'analyse approfondie, portant par conséquent sur un échantillon relativement restreint de participants.

[sed-the-69-body] – L'impact cognitif des dispositifs médiatiques sur les enfants d'âge préscolaire en situation d'apprentissage avec un adulte (BERNARD)

Extrait 4.

Ce travail représente une tentative d'explication du sens des lexèmes dans un contexte particulier. Pour y parvenir, nous avons opté pour une approche sémantico-lexicale, c'est-à-dire traiter du sens grâce à un outil d'analyse la lexicographie et une méthodologie sémantique appelée componentielle.

[lin-the-657-body] - Le sens du verbe "changer" dans le programme politique du parti écologiste "Les verts". Approche sémantique et lexicale (Cavalla)

---

## Présenter la limite d'un choix

Extrait 5.

Pour les besoins de notre étude de cas, nous avons sélectionné trois extraits de la séquence complète. Nous nous limitons à analyser la mise en route de l'activité conjointe entre l'apprenant et l'employé expert en nous focalisant sur la façon dont l'apprenant accède à la tâche ainsi que sur le régime de distribution de celle-ci. Nous montrons que la ratification de la participation de l'apprenant est construite localement et collectivement, et qu'elle a des effets au niveau de son identité telle qu'elle se construit en situation.

([sed-com-251-body] - Apprentissage, participation à l'activité et identité professionnelle (Duc) (Actes AREF 2007))

Extrait 6.

Nous avons restreint arbitrairement le critère statistique à la cooccurrence de deux mots apparaissant au moins trois fois dans le corpus, issus de trois articles différents et employés par au moins trois auteurs. Ces trois paramètres devant être simultanément remplis, ils garantissent l'authenticité des collocations retenues, en étant représentatifs des usages linguistiques de la communauté TAL.

([lin-the-276-body] - La dimension argumentative des collocations textuelles en corpus électronique spécialisé au domaine du TAL(N) (Dubreil))

## Présenter la collecte des données

Extrait 7.

Afin de recueillir les requêtes [3] formulées par les élèves, nous avons procédé à l'enregistrement des adresses web (URL[4]) des moteurs de recherche dans les historiques des navigateurs (exclusivement Internet Explorer) utilisés pendant la séance durant laquelle les élèves avaient fourni un travail autonome de recherche sur Internet. Nous avons ainsi recueilli les historiques de 9 sessions, c'est-à-dire le travail de 18 élèves, étant donné qu'ils travaillaient en binômes. À partir de ces historiques nous avons retenu 95 requêtes pour notre analyse [5].

[sed-the-356-body] - Travail collectif des enseignants et pratiques d'enseignement : le cas de la prise en charge des élèves dits en difficulté au sein de l'école primaire (Lefevre)

Extrait 8.

Cette étude exploratoire est basée sur un corpus composé de 10 textes du XIX<sup>ème</sup> collectés dans le cadre d'un projet qui porte sur l'étude des marqueurs linguistiques de la subjectivité. Ce corpus représente 830 000 mots et est essentiellement composé de romans. Nous avons également utilisé un corpus complémentaire, afin de vérifier la fréquence de certaines structures mais aussi pour bénéficier de textes plus contemporains. Il est issu de la base catégorisée de Frantext et composé de romans qui couvrent la période 1990-2005, soit 22 textes et plus de 1,5 millions de mots.

(Vannina GOOSSENS. (2006). Régularités et irrégularités dans la polysémie des noms d'affect : le cas de l'expression de la source de l'affect)

Extrait 9.

Nous avons choisi, pour un travail sur la métonymie et la synecdoque qui dépasse le cadre de cet article, de rassembler un corpus de textes de la presse généraliste sur des thèmes d'actualité susceptibles d'intéresser le grand public, ou instaurés comme tels par les journalistes, puisque ces thèmes ont été à un moment donné présentés à la « Une ». Le corpus se compose de 160 articles analysés de *Libération*, *le Monde*, *le Figaro* sur trois thèmes : la guerre du Kosovo, (corpus Kosovo ci-après) ; les accords conduisant à la participation au gouvernement du parti d'extrême-droite (FPÖ) en Autriche (corpus Autriche ci-après) ; le report du Congrès pour le vote sur le Conseil Supérieur de la Magistrature (corpus CSM ci-après).

Notre but en choisissant la presse était de rencontrer des figures présumées non littéraires, et donc de pouvoir considérer sous un angle différent ce qui est appelé généralement « figure de style ».

(Michelle Lecolle, Figures et références plurielle en corpus journalistique)

### III. Exercice lacunaire sur les verbes synonymes (15m)

Activité : Dans cette étape, les apprenants travaillent en groupe et doivent choisir le bon verbe selon le sens donné. Il s'agit de groupes de verbes synonymes donnés qui servent à présenter le choix de l'approche, la limite d'un choix et la collecte des données. L'objectif est d'amener l'apprenant à distinguer les verbes synonymes dans un contexte donné et à s'approprier le sens des verbes.

Explication : L'apprenant entre dans la phase d'appropriation des éléments linguistiques. Les verbes abordés dans la phase d'observation présentent peu de variations au niveau syntaxique, mais ils présentent des caractéristiques particulières au niveau sémantique. Cette étape permet à l'apprenant de mieux saisir la différence de sens dans un groupe de verbes synonymes et ensuite de mieux utiliser les verbes dans un contexte approprié.

Consigne : À l'aide de l'outil CNTRL en ligne : <http://www.cnrtl.fr/synonymie/>, choisissez le verbe qui convient le mieux dans la phrase selon le sens donné entre parenthèses. Conjuguez le verbe si besoin.

- 
1. Nous avons isolé plus de 200 objets économiques dans notre corpus, que nous avons **regroupés** en classes fonctionnelles. (classer les éléments dans un groupe selon un critère cohérent)
  2. Enfin, la démarche d'évaluation fût complétée en **recueillant** des données d'observation sur la qualité de l'environnement physique de la communauté (propreté, sécurité routière, qualité des bâtiments et des infrastructures, etc.). (chercher et assembler des données en mettant à part comme information pour étude)
  3. Nous avons également **rassemblé** un corpus parallèle, constitué des versions françaises et anglaises de rapports du parlement européen, disponibles en ligne sur le serveur d'Europarl 144. (mettre ensemble des données, des documents en vue d'une exploitation)
  4. Poudat qui semble plus restrictive, dans le sens où elle entend s'appuyer sur des revues notoires, plus diffusées dans le champ, alors que nous avons **retenu** aussi bien des revues nationales que des revues locales. (Conserver après un examen critique, après une sélection.)
  5. Or, la qualité et la quantité de données récoltées se sont avérées si pauvres, que nous avons **adopté** en définitive une démarche inductive inverse. (choisir de propos délibéré quelque chose pour le faire sien, s'y rallier ou être d'accord avec)
  6. C'est pourquoi nous **privilegions** l'observation ethnographique et les enregistrements audio-vidéo comme moyens de constitution de nos données. (donner une valeur particulière à quelque chose)
  7. Étant donné la variété des situations, on **se limitera à** l'emploi en français, au singulier, des articles définis (le, la, l'), indéfinis (un, une), et du nom sans article. Cela paraît bien sûr très restrictif, mais nous verrons que la variété des situations demande de commencer à un niveau modeste. (se fixer des limites et se contenter de quelque chose, les autres possibilités sont interdites)
  8. Si l'on **se restreint**, pour simplifier, à l'évaluation phrase par phrase, on constate que l'ensemble des traductions acceptables d'une phrase donnée est potentiellement vaste (selon le domaine et la longueur) mais borné. (se réduire à quelque chose même si d'autres choix sont possibles)

#### **IV. Production écrite à partir des éléments liés à la méthodologie (15m)**

Activité : Dans cette étape, l'apprenant travaille individuellement et doit rédiger un paragraphe sur la méthodologie à partir des éléments donnés. Il doit utiliser les constructions étudiées.

Explication : Il s'agit toujours de la phrase d'appropriation des éléments linguistiques.

Consigne : À partir des éléments fournis ci-dessous, constituez en quelques lignes la méthodologie adoptée et la collecte des données à l'aide des constructions étudiées.

Les éléments de la méthodologie :

- un corpus large de données orales et écrites en allemand et en français
- des questionnaires et des entretiens semi-directifs conduits avec les élèves dans les trois lycées
- En raison d'espace limitée, l'analyse des données orales de deux élèves des lycées de Buc et de Sarrebruck seulement

## V. Classement des constructions étudiées (10m)

Activité : L'apprenant doit faire un récapitulatif des différentes constructions étudiées.

Explication : Cette étape vise à fixer les acquis de l'apprenant et lui permet de faire un récapitulatif. En plus, il peut approfondir ses connaissances sur les différents sens des constructions.

Consigne : Vous venez d'apprendre des expressions servant à présenter un choix de l'approche, la limite d'un choix et la collecte des données. Faites un récapitulatif des expressions étudiées et présentées ci-dessous :

1. Nous nous limitons à étudier.../à l'étude
2. Rassembler des articles/des textes
3. Nous adoptons/retenons/privilégions l'approche
4. Regrouper les données
5. Nous choisissons de faire
6. On se restreint à l'analyse...
7. Nous optons pour la méthode
8. Les données sont recueillies

Présenter un choix favorable	Présenter la limite d'un choix	Présenter la collecte des données
Nous adoptons l'approche...	On se limite à	Les données sont recueillies...

## VI. Production écrite d'un paragraphe sur la méthodologie de recherche (20m)

Activité : L'apprenant doit répondre à des questions sur la méthodologie de recherche. Dans un deuxième temps, ses réponses lui permettent de présenter en quelques lignes sa méthodologie de recherche. Des expressions déclencheurs sont données pour faciliter la rédaction.

Explication : Cette étape vise à consolider les connaissances acquises de l'apprenant.

Consigne : Dans le cadre de votre mémoire, vous allez présenter en quelques lignes l'approche adoptée et la collecte des données en essayant de répondre aux questions suivantes :

1. Avec quel objectif de recherche, souhaitez-vous constituer un corpus ?
2. Quel est le type de corpus/de données recueillies (articles, productions écrites/orales, etc.)?
3. De quelle façon les données ont-elles été recueillies (enregistrement, entretien, questionnaire, etc.) ? Au près de quel public ?
4. Concernant la représentativité du corpus, y a-t-il une limite de recherche ?

Pour commencer, on vous propose plusieurs expressions :

*Ayant pour objectif de..., nous avons adopté l'approche ...*

*L'étude est basée sur l'approche/le corpus...*

*Pour mener à bien cet objectif, nous nous sommes appuyés sur...*

*Afin de..., nous avons choisi de....*

## **Fiche pédagogique de la quatrième séance (120m): l'illustration et le constat**

### **I. Découverte du lexique et des constructions qui portent sur**

#### **l'illustration et le constat (20m)**

Activité : Dans cette étape, à partir d'un document écrit, les apprenants travaillent en binômes pour repérer le lexique et les constructions qui portent sur l'illustration et le constat en répondant aux questions.

Explication : Il s'agit d'une étape de découverte qui permet à l'apprenant de comprendre et de se familiariser avec la façon dont on éclaire son propos par une illustration (constat et interprétation).

Consigne : Lisez le document suivant et répondez aux questions :

<b>Valeurs</b>	<b>« Ne pas aimer » et « Ne pas aimer du tout »</b>	<b>« Aimer » et « Aimer bien »</b>
Lecture	15,91%	60,57%
Vocabulaire	21,70%	48,59%
Grammaire	22,86%	54,24%
Conjugaison	25,86%	44,94%
Résumé	64,28%	19,37%
Discussion	50,40%	29,38%

**Tableau 1 - Investissement des sous-catégories de la discipline de français**

Le tableau montre que plus de la moitié des élèves n'aiment pas le résumé et la discussion. On peut penser que ce désintérêt tient, entre autres, à deux raisons principales. La première pourrait être d'ordre générique. Nous faisons l'hypothèse en effet que les élèves ne maîtrisent pas, voire ne disposent pas de représentations précises des règles de codification et de structuration non linguistiques des textes à produire (Rastier, 1989), en rapport avec les tâches d'écriture demandées. Cette hypothèse peut être posée dans la mesure où 44% des élèves interrogés considèrent, en réponse à une autre question (cf. 4.3.4), qu'il faut, pour écrire, avoir des modèles en tête. La deuxième relèverait de « l'image scripturale » dévalorisante qu'ils se font de leurs textes par rapport aux textes des auteurs de la littérature qu'ils étudient à l'école (Delcambre & Reuter, 2002). Se mesurant souvent à l'aune de ces derniers, ils ont construit une « image de soi » comme de piètres scribes. En effet, quand il s'agit d'attribuer la valeur de « véritable écrit », les écrits qu'ils produisent en milieu scolaire ne sont majoritairement pas qualifiés comme de véritables écrits, contrairement à ceux d'écrivains et de journalistes (cf. 4.3.2). Par ailleurs, nous observons un écart de 25% entre l'investissement vis - à - vis des outils de la langue et vis - à - vis de l'écriture. Les élèves préfèrent les exercices de grammaire aux travaux d'écriture. Étant donné cet écart, on peut se demander s'il n'y a pas, dans leurs conceptions sur les modalités d'acquisition de l'écriture, une certaine dissonance entre compétence linguistique et savoir écrire.

Ali Djaroun, Quand les lycéens algériens parlent de leurs compétences scripturales : enjeux et perspectives didactiques.

1. Qu'est-ce que l'auteur a observé dans le tableau ? Soulignez l'élément linguistique qui justifie votre choix.
2. Comment l'auteur a-t-il interprété ce constat ? Soulignez l'élément linguistique qui justifie votre choix.
3. Trouvez des mots synonymes qui peuvent remplacer le verbe « observer » dans le texte à l'aide de l'outil CNTRL en ligne : <http://www.cnrtl.fr/synonymie/> (ex : *remarquer, constater, noter, voir ?*)

## II. Analyse des phénomènes linguistiques liés à l'illustration

### et au constat (40m)

Activité : Cette étape est divisée en deux champs sémantiques : l'illustration et le constat. Les apprenants travaillent en groupe et seront amenés à analyser des extraits courts pour repérer des constructions liées à ces deux champs sémantiques.

Explication : Il s'agit d'une étape intermédiaire qui aide l'apprenant à faire une progression propice à la compréhension des phénomènes linguistiques nouveaux par rapport à l'illustration et au constat dans la partie de l'analyse.

Consigne : Relevez les expressions qui portent sur l'illustration. Relevez les expressions qui portent sur le constat.

---

## Éclairer le propos par une illustration graphique

### Extrait 1.

L'utilisation du langage parlé implique un certain niveau de dialogue qu'il importe d'intégrer dans le système. Peu de systèmes prennent effectivement en compte cet aspect du problème mais c'est une voie de recherche importante. Au total, le processus de compréhension est complexe et multiforme. La figure 16 en illustre les différentes composantes, depuis le décodage acoustico-phonétique jusqu'au dialogue. Les sources de connaissances impliquées sont précisées.

(J.P. Haton & M.C. Haton, 1989, L'intelligence artificielle, Paris, PUF : 97)

### Extrait 2.

Or le tableau 3 nous indique la nette prédominance de deux catégories de faits, les violences et les dévoiements de la fonction policière. Nous avons dit plus haut que les violences représentaient 16,7 % des sanctions prises pour la période 1996-2001, nous retrouvons des chiffres analogues ici puisque, pour la période 1993-1999, 18,1 % des faits pour lesquels les policiers sont traduits devant un conseil de discipline sont des cas de violences.

([spo-art-380-body] - " Violences illégitimes et publicité de l'action policière (Moreau de Bellaing ) (Politix 87))

### Extrait 3.

Les résultats du Tableau 4 montrent qu'à 5 ans les enfants utilisent principalement un seul indice pour marquer l'émotion dans leurs dessins, tandis qu'à partir de 8 ans, deux ou trois indices marquent l'émotion dans les dessins expressifs. Par ailleurs, chez les enfants de 11 ans, il apparaît que la tendance à utiliser trois indices ou plus pour marquer graphiquement l'émotion est plus prononcée que chez les enfants de 8 ans et les adultes. Le test du X<sup>2</sup> (avec un seuil alpha de ,05) révèle que, toute émotion confondue, les dessins expressifs à un seul indice sont significativement plus nombreux chez les enfants de 5 ans (12/13) que chez les enfants de 8 ans (9/33), X<sup>2</sup> (1) = 13,39, p - ,01.

([psy-art-286-body] - Expression des émotions dans le dessin d'un homme chez l'enfant de 5 à 11 ans (Brechet, Picard, Baldy) (Canadian journal of experimental psychology))

### Extrait 4.

Ce phénomène d'ébranlement des certitudes a été expliqué par Nonnon (1996) en référence au fait que dans une situation où des connaissances nouvelles sont, a priori, à



élaborer, les élèves ne sont pas en mesure d'exprimer des points de vue et des opinions stables ou figés. L'argumentation dialoguée peut impliquer un travail cognitif-langagier accru, avec des tentatives d'association ou de dissociation des notions, **comme le montre l'extrait** tiré du corpus reproduit dans la Figure 13 ci-dessous (où aucun changement d'opinions n'est exprimé).

([psy-hdr-647-body] - Recherches sur l'élaboration de connaissances dans le dialogue (Baker))

## Interprétation à partir d'un constat

Extrait 5.

Plus grand est ce nombre, plus fort est l'effet. Par exemple, **dans le tableau D** (résultats pour Schaerbeek), **on peut constater que** l'origine exerce un impact plus important sur le comportement électoral que les trois autres variables structurelles, à savoir le genre (dont l'effet n'est pas significatif), le niveau d'éducation et la position socio-économique. **Ces résultats montrent que** les comportements électoraux des groupes ethniques sont en partie indépendants de leur niveau d'éducation ou de leur statut socioéconomique.

([spo-art-153-body] - Les comportements électoraux des minorités ethniques à Bruxelles (Rea, Jacobs, Teney, Delwit) (Revue française de science politique Vol. 60))

Extrait 6.

**L'observation de la figure 1 montre que**, quelle que soit la tranche d'âge, les enfants issus du milieu cadre réalisent davantage de liaisons facultatives conformément à la cible adulte que les enfants issus du milieu ouvrier. **Nous notons d'ailleurs un effet significatif** du milieu social (ANOVA 5,  $F(1,179) = 11,123$ ,  $p = 0,001$ ). **Nous remarquons également qu'**à 2 - 3 ans, les différences entre les enfants des deux milieux sont peu importantes et qu'elles s'accroissent progressivement durant le développement. À 5 - 6 ans, la différence entre les taux de réalisation de liaisons facultatives chez les enfants des 2 milieux (40,70% chez les enfants de parents cadres et 21% chez les enfants de parents ouvriers) devient significative (Holm - Sidak,  $p = 0,011$ ).

(Aurélié NARDY, Nathaël MARTIN, Usages langagières enfantins et variables sociolinguistiques : convergences microsociologiques et stratification macrosociologique.)

Extrait 7.

**Les exemples du tableau 2**, concernant les sports collectifs, **mettent en évidence** les transformations didactiques opérées sur ces situations motrices par les enseignants.

On s'aperçoit ici que la transformation d'un sport en un jeu didactique s'étend sur une échelle allant de la dénaturation à la modification légère du jeu sportif de référence. Dans le cas du volleyball, nous observons souvent que sa pratique s'accomplit dans un 3 contre 3 ou encore au sein d'un 4 contre 4. Cette situation est rencontrée fréquemment au lycée car l'évaluation certificative du baccalauréat est effectuée dans ces conditions.

([sed-art-52-body] - Des pratiques sociales aux pratiques scolaires en EPS : des logiques de scolarisation plurielles (Dugas) (Revue Française de Pédagogie))

Extrait 8.

Conformément aux résultats obtenus par Elliot et McGregor (2001) pour la version en langue anglaise du questionnaire, l'analyse met en évidence l'existence de quatre facteurs distincts qui expliquent en tout 74% de la variance. Comme on peut le voir dans le tableau 7, le premier facteur comprend les trois items de performance-approche et explique 30.9% de la variance. Le second facteur explique 16.6% de la variance et correspond aux trois items de maîtrise-approche. Le troisième facteur explique pour sa part 14.7% de la variance. Il correspond au but de maîtrise-évitement et comprend les trois items y correspondant. Il est intéressant de noter que les items de maîtrise-approche sont très peu représentés sur le facteur 2 et réciproquement. Enfin, le dernier facteur regroupe les trois items d'évitement de la performance et explique 12.1% de la variance. Tout comme pour les items de maîtrise, ces trois items sont très peu représentés par le facteur 1 (performance-approche) et réciproquement [43].

([psy-the-504-body] - Conflit sociocognitif et buts d'accomplissement : effets interactifs sur l'apprentissage et le mode de régulation du conflit (Darnon))

### III. Exercice lacunaires sur les verbes de constat et les verbes d'illustration (15m)

Activité : Les apprenants travaillent en groupe et doivent choisir parmi plusieurs verbes proposés, le (les) verbe (s) qui conviennent le mieux dans la phrase. En outre, il doivent rayer le verbe inapproprié. Il s'agit de verbes d'illustration et de constat qu'on a travaillé dans la phase précédente. Pour faciliter la tâche, ils peuvent consulter le dictionnaire TLFi en ligne.

Explication : L'apprenant entre dans la phase d'appropriation des éléments linguistiques. Après l'observation des extraits dans la phase précédente, ici, l'objectif est d'amener l'apprenant à travailler sur les différentes constructions des verbes et à distinguer les verbes synonymes.

Consigne : À l'aide du dictionnaire TLFi, rayez le mauvais verbe ou les mauvais verbes qui ne peut (peuvent) pas entrer dans la phrase et choisissez le bon verbe ou les bons verbes. Conjuguez le verbe si besoin.

1. Dans le tableau ci-dessous \_\_\_\_\_ (montrer, indiquer, présenter) les acteurs ayant été choisis régulièrement comme destinataires des interrelations des enseignants (rangs 1

---

ou 2 dans plus de 50% des cas). Plusieurs faits peuvent \_\_\_\_\_ (voir, remarquer, constater).

2. Nous avons cherché à \_\_\_\_\_ (observer, remarquer, constater) les effets du sexe, de l'âge et de la CSP sur Rn. Comme le \_\_\_\_\_ (signaler, montrer, illustrer) la figure 1, deux phénomènes sont à \_\_\_\_\_ (remarquer, noter, voir).

3. Si nous comparons les réponses à la question ouverte des enfants issus de milieux défavorisés fréquentant des écoles socialement mixtes et ceux fréquentant des écoles non mixtes, nous \_\_\_\_\_ (observer, noter, constater) que les enfants faisant spontanément référence à l'interaction sont significativement plus nombreux dans les écoles mixtes ( $\chi^2 = 4,458$  à 1ddl;  $p = 0,034$ ).

4. Le tableau suivant \_\_\_\_\_ (présenter, révéler, témoigner) de la présence féminine dans les agrégations d'anglais et d'allemand entre 1883, date de l'ouverture du concours aux femmes, et 1944.

5. Le tableau 1, proposé par Hu et Hu (2008), \_\_\_\_\_ (montrer, indiquer, révéler) bien la différence entre les études portant sur la communication de masse et les autres champs d'étude en communication, et la progression de la communication organisationnelle ces dernières années. Le tableau permet de \_\_\_\_\_ (voir, constater, noter) d'ailleurs que l'étude de la communication interpersonnelle a commencé à se développer au même moment que la communication organisationnelle, et que les études sur la communication de groupe se développent tout juste.

6. Ce regroupement permettra un traitement unitaire dans les textes de ces classes de FL à l'aide d'un transducteur d'états finis manipulant les enregistrements (Cf. Figure 1). Une table correspondant à un paradigme de FL regroupant les noms, leurs collocatifs et leurs propriétés est ainsi facilement créée comme on le \_\_\_\_\_ (constater, voir, remarquer) dans le tableau 2 ci-dessous. Le modèle \_\_\_\_\_ (montrer, révéler, illustrer) dans la figure 1.

## **IV. Exercice sur les adverbes associés aux verbes de constat et au verbe d'illustration (15m)**

Activité : Dans cette étape, les apprenants travaillent en groupe et doivent choisir et insérer les adverbes proposés dans la phrase. Il s'agit des adverbes « localisateurs », du type de *ici*, *ci-dessous* et des adverbes de « relation logique » du type de *cependant*, *également*, *donc*, *d'ailleurs*. Pour chaque phrase proposée, plusieurs choix sont possibles, l'objectif est que la phrase soit mieux structurée avec les adverbes insérés aux bons endroits.

Explication : L'apprenant est toujours dans la phase d'appropriation des éléments linguistiques. Les adverbes de « localisation » et de « relation logique » contraignent l'emplacement de la figure et côtoient les verbes de l'illustration et du constat. Comme les adverbes de « relation logique » sont très fréquents dans l'écrit scientifique, la première analyse des erreurs d'apprenants nous conduit à les faire travailler sur la distribution des adverbes dans la phrase. Cette étape permet à l'apprenant de mieux

---

cerner l'utilisation de ces adverbes et donc – nous l'espérons – de mieux les comprendre en termes de distribution syntaxique.

Consigne : Remettez l'adverbe au bon endroit du texte.

### **ci-dessous, ainsi, donc**

1. Le Tableau 3 détaille ces paradigmes verbaux. Nous constatons que être et avoir présentent cinq formes différentes sur les six du paradigme. <...> Cette première analyse permet de constater que le système verbal enfantin porte beaucoup moins bien l'étiquette de « pauvreté morphologique » que son équivalent adulte...(Katérina PALASIS-JOURDAN, syntaxe, acquisition et dialectologie : la pluridisciplinarité appliqué à l'étude du sujet nul enfantin)

### **enfin, ci-après, cependant**

2. Le tableau montre que la forme allemande werden + infinitif, très contestée comme Temps (entre autres à cause de sa forme composée), a toutes les caractéristiques formelles d'un Temps futur (cf. Rebotier, 2006). On ne constate pas de véritable différence entre les Parfaits allemand et français, alors qu'on attendait une grammaticalisation plus avancée en français à cause de son emploi généralisé à l'oral. (Aude Bebotier, Les critères de grammaticalisation pour les temps composés en français et en allemand)

### **également, voici enfin**

3. Ces tableaux permettent de mettre en évidence une combinaison privilégiée entre les objets sales et rate avec le prédicat grow, plus qu'avec tout autre prédicat. Nous observons une forte association entre réduire et coûts/taux en français, alors que l'anglais utilise indifféremment to reduce ou to cut costs/rates. La présentation des combinaisons les plus fréquentes avec l'objet conceptuel CURRENCY. ([lin-the-614-body] - Contribution à la traduction de documents économiques anglais : importance de l'extraction et de la représentation de sens (Ledouble))

### **ici, , d'ailleurs, même**

4. On trouve des différences entre la forme la plus souvent produite par les apprenants et la forme la plus fréquente dans l'input, on peut voir qu'il y a une assez bonne correspondance pour les premiers et les derniers verbes de la liste. On peut noter que les apprenants ont uniquement produit le verbe préférer à la forme courte du présent, ce qui correspond à l'hypothèse d'un lien préférentiel entre les verbes d'état et la forme courte du présent. (Formes par défaut et fréquence dans l'input en FLE, Anita Thomas)

## **V. Classement sémantique (10m)**

Activité : L'apprenant doit faire un récapitulatif des différentes constructions étudiées.

Explication : Cette étape vise à fixer les acquis de l'apprenant et lui permet de faire un récapitulatif. De plus, il peut approfondir ses connaissances sur les différents sens des constructions.

Consigne : Vous venez d'apprendre des expressions liées à l'illustration et au constat. Faites un récapitulatif des expressions étudiées.

Parler d'une illustration	Parler d'un constat	Proposer une interprétation
Le tableau ci-dessous nous indique que...	On constate que...	Il semble que...

## VI. Production écrite d'un paragraphe à partir de l'observation d'une figure (20m)

Activité : L'apprenant doit observer et lire les informations indiquées dans l'encadré et sur la figure ci-dessous.

Dans un deuxième temps, il doit présenter en quelques lignes ce que le graphique illustre en utilisant les expressions étudiées sur l'illustration et le constat.

Explication : Cette étape vise à consolider et vérifier les connaissances acquises de l'apprenant.

Consigne: À partir éléments fournis ci-dessous, observez la figure 1 et constituez en quelques lignes une analyse/interprétation basée sur l'observation de la figure.

Objectif : constater la présence de troubles langagiers chez les ESML

Illustration : Les domaines langagiers déficitaires, mis en évidence par les différentes épreuves du BILO

Public : deux groupes d'enfants (les ESML 6-9 ans et 9-11 ans)

Constat : 1. Des troubles langagiers prédominant en morphosyntaxe et en phonologie (déficits les plus cités dans la littérature) ; 2. les troubles du vocabulaire et de la phonologie diminuent davantage avec l'âge que les troubles de la morphosyntaxe qui seraient donc plus persistant.

([lin-the-505-introduction] - Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage (Dugua))

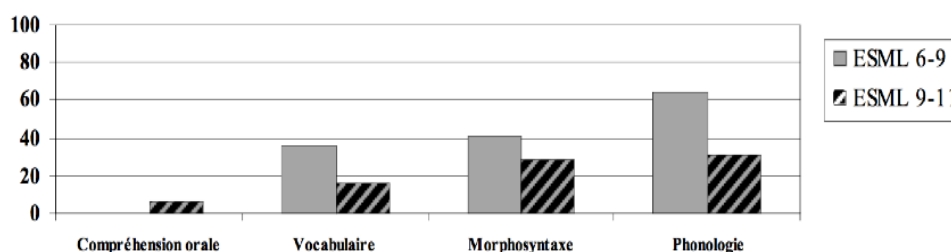


Figure 1 - Domaines langagiers déficitaires : % de scores < -1,65 ET

---

## Fiche pédagogique de la cinquième séance (120m) : faire un rappel de l'étude présentée et résumer les résultats obtenus

### I. Découverte du lexique et des constructions qui visent à faire un rappel de l'étude présentée (40m)

Activité : Dans cette étape, à partir de quatre documents écrits, les apprenants travaillent à deux pour distinguer l'introduction de la conclusion à travers des verbes (en particulier le temps des verbes, et les différentes collocations liées à l'hypothèse) en repérant les indices linguistiques qui caractérisent la conclusion.

Comme l'introduction est déjà abordée dans la première séance, l'apprenant sera conduit à travailler sur l'introduction et puis sur la conclusion pour comparer les verbes à aborder dans ces deux parties d'un mémoire. A la fin de l'exercice, il est proposé de trouver des synonymes pour les collocations associées à l'hypothèse à l'aide de l'outil CNTRL. La mise en commun pour cet exercice ne sera faite qu'après la présentation de l'utilisation du corpus (Scientext).

Explication : Il s'agit d'une étape de découverte qui permet à l'apprenant de comprendre que les verbes d'examen dans la conclusion sont souvent au temps passé, cela s'explique par le fait que dans cette partie, l'auteur fait un rappel de l'étude présentée. Et cela constitue une caractéristique importante permettant de distinguer l'introduction de la conclusion.

Ensuite, concernant les deux documents sur l'hypothèse, l'apprenant va découvrir les collocations associées à l'hypothèse qui apparaissent dans des parties textuelles différentes. D'ailleurs, l'exercice sur les synonymes des collocations ne sera corrigé qu'après la présentation de l'utilisation du corpus, ce qui leur permettra de découvrir l'avantage du corpus.

Consigne : Vous travaillez à deux. Lisez, comparez les deux documents suivants et répondez aux questions.

#### Document 1

En fait, l'effet d'injure est le résultat de l'interaction de plusieurs composantes, dont certaines proprement linguistiques, d'autres extra- ou paralinguistiques.

Dans ce qui suit, nous allons analyser le rôle de chacune de ces composantes et surtout une des constructions syntactico-énonciatives qui sert l'injure, à savoir l'apostrophe, considérée dans deux de ses réalisations : comme apostrophe neutre vs injurieuse.

À travers une comparaison d'apostrophes construites sur un N titre, sur un N insultant ou hypocoristique ainsi que sur des N dénotatifs modifiés ou non, nous étudierons tant le fonctionnement d'un type particulier de N (les N insultants) que les contraintes sémantiques qui pèsent sur la construction même de l'apostrophe.

([lin-com-518-introduction] - Pour une étude de l'apostrophe injurieuse (MATEIU)  
(Autour des langues et du langage : perspective pluridisciplinaire))

#### Document 2

Au cours de cette contribution, nous avons tenté d'établir un lien entre la théorie économique du consommateur et le concept de développement soutenable. Pour ce faire, nous avons procédé en deux étapes. Dans un premier temps, nous proposons un aperçu de la consommation dans la théorie économique, en partant de l'analyse économique standard pour mettre en avant ses limites : préférences données, absence d'explication structurelle des processus de consommation. Nous avons ensuite analysé l'apport de la "nouvelle économie du consommateur" qui dégage des processus de consommation en critiquant la théorie économique traditionnelle. Cependant, au terme de la première partie, il apparaît qu'il n'existe pas de théorie spécifique de la Critiquer/théorie 109 consommation : le producteur et le consommateur sont traités de façon symétrique d'un point de vue analytique, sans prendre en compte la dimension sociale de la consommation.

(Patrick Jolivet, Audrey Aknin Consommation et développement durable : définition d'une problématique)

1. Quels sont les sujets abordés dans les deux documents ci-dessus ?
2. Les deux documents appartiennent respectivement à quelle partie de l'article (introduction, théorie, méthodologie, conclusion) ? Quels sont les indices linguistiques qui justifient votre choix ?
3. Comparez les indices linguistiques repérés dans ces deux documents, quelles sont les différences ? Quelles sont les fonctions typiques de ces indices linguistiques ? (ou à quoi servent ces indices linguistiques ?)

Consigne : Vous travaillez à deux. Lisez, comparez les deux documents suivants et répondez aux questions.

#### Document 3

Les noms d'affect (désormais N\_affect) tels *peur*, *colère*, *joie*, *amour*, etc., sont sémantiquement complexes et, de plus, fortement polysémiques. Nous nous proposons d'étudier les caractéristiques de cette polysémie dans le cadre d'un traitement unitaire des N\_affect permettant de poursuivre l'entreprise typologique initiée sur ces noms à partir de leurs propriétés combinatoires (Tutin *et al.*, 2006, Goossens, 2005). Nous nous demanderons en particulier si ces noms obéissent à des modèles de polysémie spécifique, et sur quelles bases on peut fonder les régularités observées. Nous émettons en effet l'hypothèse qu'il existe des schémas de polysémie régulière qui pourraient s'avérer caractéristiques de cette classe de noms et susceptibles de nous renseigner sur leurs caractéristiques sémantiques. Afin de tester cette hypothèse, nous avons décidé de nous intéresser dans un premier temps à un seul type de dérivation polysémique régulière, les cas où le N\_affect semble exprimer en réalité la source de cet affect.

([lin-com-403-introduction] - Régularités et irrégularités dans la polysémie des noms d'affect : le cas de l'expression de la source de l'affect (Goosens) (Autour des langues et du langage: perspective pluridisciplinaire))

#### Document 4

Les estimations de notre modèle semblent valider l'interprétation de la relation entre taux du crédit et taux de marché en terme de « facturation du coût des ressources bancaires ». Elles justifient, en particulier, l'hypothèse d'existence d'une structure de financement et de refinancement des crédits nouveaux <...>. Par ailleurs, nos résultats d'estimation semblent conforter l'hypothèse de concurrence imparfaite sur le marché du crédit. Dans ce cadre, l'analyse des estimations conduit à une élasticité de la demande globale de crédit de l'ordre de 0.7 et à la conclusion que, sur le marché du crédit, les « 3 grandes banques » sont soumises à une plus forte concurrence que les banques de plus petite taille.

([eco-art-443-conclusion] - La relation entre le taux des crédits et le coût des ressources bancaires Modélisation et estimation sur données individuelles de banques (Laurent, Sevestre) (Annales d'économie et de statistique, 59, 199-226, 2000))

1. Quels sont les sujets abordés dans les deux documents ci-dessus ?
2. Les deux documents appartiennent respectivement à quelle partie de l'article (introduction, théorie, méthodologie, conclusion) ? Quels sont les indices linguistiques qui justifient votre choix ?
3. Observez le nom « hypothèse » dans ces deux documents. Quels sont les différents verbes associés ? Peut-on les remplacer par d'autres verbes ? (Vous pouvez consulter le dictionnaire TLFi. A la fin de la séance, vous pouvez tester votre proposition avec le corpus.

## II. Analyse des phénomènes linguistiques liés aux résultats

**(40m)**

Activité : Cette étape est divisée en deux champs sémantiques : faire un rappel du travail présenté (*Nous avons analysé/étudié/réalisé...*) et résumer l'essentiel des résultats. Le dernier est divisé en deux champs : faire un retour sur l'hypothèse (*l'étude permet de vérifier notre hypothèse*) et extraire des résultats (*le résultat se dégage de, on peut conclure à..., etc.*). Les apprenants travaillent en groupe et seront amenés à analyser des extraits courts pour repérer des constructions liées à ces deux champs sémantiques.

Explication : Il s'agit d'une étape intermédiaire qui aide l'apprenant à faire une progression propice à la compréhension des phénomènes linguistiques nouveaux par rapport à la présentation des résultats dans la conclusion.



---

Consigne : Relevez les expressions qui permettent de faire un rappel du travail présenté.

Extrait 1.

**Nous avons étudié** le jeu stratégique de marché avec bien-monnaie dans le modèle à générations imbriquées. La monnaie est considérée comme un bien durable sans utilité. De ce fait, c'est un modèle proche de celui de Samuelson et il permet une analyse claire des relations entre l'équilibre de concurrence imparfaite et celui de concurrence parfaite.

L'analyse des jeux stratégiques de marché dans le modèle à générations imbriquées doit se poursuivre en considérant une monnaie-signal c'est-à-dire un modèle de création monétaire.

([eco-art-39-conclusion] - Équilibres monétaires du jeu stratégique de marché dans le modèle à générations imbriquées (Morogues Francis) (Annales d'économie et de statistique, 54, 1-21, 1999))

Extrait 2.

Dans cet article, **nous avons analysé** par quels mécanismes le niveau des allocations chômage peut affecter le taux de chômage de long terme dans le cadre d'un modèle de négociations salariales. Pour ce faire, **nous avons explicitement distingué** une allocation d'assurance chômage, qui verse pour une durée limitée une allocation dont le montant est en partie indexée sur le dernier salaire reçu, et d'une allocation d'assistance chômage forfaitaire et illimitée.

**Nous avons alors montré qu'**il y avait deux canaux essentiels par lesquels le niveau des allocations chômage affecte le taux de chômage de long terme.

([eco-art-308-conclusion] – L'impact de l'assurance chômage et de l'assistance chômage sur le chômage d'équilibre (Etienne) (Annales d'économie et de statistique, 53, 31-41, 1999))

Extrait 3.

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. Pour ancrer les aspects théoriques associés à ces trois concepts, **une analyse a été réalisée à partir de données recueillies** dans le cadre d'un projet portant sur le TEC, entre collègues qui interviennent auprès d'élèves d'un même cycle d'apprentissage (Lafortune, 2004). La synthèse des résultats de cette analyse comporte les points suivants.

[sed-com-109-conclusion] - Dimension affective, travail en équipe de collègues et bien-être des personnels scolaires (Lafortune) (Actes de AREF)

Consigne : Relevez les expressions qui permettent de résumer l'essentiel des résultats obtenus en faisant un retour sur l'hypothèse.

Extrait 4.

Nos analyses typologiques nous ont conduits à postuler un certain nombre de tendances structurelles et organisationnelles qui **confirment, prolongent ou remettent en cause des résultats d'études** antérieures sur la syllabe (chapitre III). **Après avoir établi une typologie** des langues en fonction de la structuration syllabique des unités lexicales (nombre de syllabes, de phonèmes,...), nous nous sommes intéressés à l'organisation interne de la syllabe. La mise en évidence de structures syllabiques extrêmement favorisées, ainsi que de tendances nettes quant à l'apparition des différentes structures dans les langues a **conforté l'hypothèse** de l'existence de contraintes pesant sur l'organisation syllabique.

(Isabelle Rousset. Structures syllabiques et lexicales des langues du monde données, typologies, tendances universelles et contraintes substantielles. Linguistics. Université Stendhal – Grenoble III, 2004. French.)

Consigne : Relevez les expressions qui permettent de résumer l'essentiel des résultats.

Extrait 5.

Ces quelques éléments d'analyse présentés dans cette contribution ne permettent certes pas de **conclure à** l'existence d'effets d'une raison numérique à l'œuvre. Ils restent insuffisants pour rendre intelligible la situation relativement inédite que nous traversons et qui voit la transformation des actes de lecture, des procédés graphiques et des supports de stockage des écrits. **Ces résultats partiels permettent** néanmoins, à titre exploratoire, de **dégager quelques pistes** relatives aux évolutions du rapport à l'écrit dans les environnements numériques.

Frédérique BROS, De la raison graphique à la raison numérique ? Evolutions du rapport à l'écrit dans les espaces virtuels

Extrait 6.

**Deux résultats principaux se dégagent de** cette étude IRMf : (i) concernant la représentation inter-hémisphérique, les deux tâches choisies induisent une latéralisation prégnante et (ii) concernant la représentation intra-hémisphérique, ces tâches permettent de mettre en évidence les aspects sémantiques et phonologiques du langage classiquement décrits dans la littérature.

([lin-com-texte\_communication\_scientext\_52-conclusion] - Évaluation de la représentation intra- et inter-hémisphérique du langage chez les sujets sains Étude comportementale en champ visuel divisé et étude IRMf (Autour des langues et du langage : perspective pluridisciplinaire))

### III. Classement sémantique (10m)

Activité : L'apprenant doit faire un récapitulatif des différentes constructions étudiées.

Explication : Cette étape vise à fixer les acquis de l'apprenant et lui permet de faire un récapitulatif. De plus, il peut approfondir ses connaissances sur les différents sens des constructions.

Consigne : Classez les différentes expressions étudiées dans le tableau ci-dessous.

**Tableau : Formes et fonctions de la conclusion**

Marques linguistiques	Exemples	fonction
Localisation dans le texte		
Temps et modes grammaticaux		
Expressions liées à l'étude présentée		
Expressions liées au « résultat »		
Expressions liées à « l'hypothèse »		

#### **IV. Introduction à l'utilisation du corpus (Scientext et dicorpus) : présentation du corpus, faire des requêtes (30m)**

Activité : Dans un premier temps, on va présenter le corpus Scientext (le choix des textes et la requête libre). En reprenant l'exercice sur les collocatifs du mot « hypothèse », on montre comment faire une recherche sur tous les verbes ayant comme objet direct le mot « hypothèse ». Cette présentation brève permet à l'apprenant de pouvoir accéder aux exemples authentiques selon leur besoin dans leur future rédaction. Dans un deuxième temps, on présente l'outil Dicorpus avec des exemples présélectionnés. Les deux entrées (par expression et par sens) sont montrées.

Explication : Il s'agit d'une présentation du corpus et de l'outil que l'apprenant pourrait utiliser plus tard. L'objectif est d'amener l'apprenant vers l'autonomie soit en faisant leur propre recherche dans le corpus soit en consultant la ressource de Dicorpus.

---

## Fiche pédagogique de la sixième séance (120m) : vers l'autonomie de rédaction à l'aide de Dicorpus

### I. Rappel de la présentation de corpus Scientext et de l'outil

#### Dicorpus. (20m)

Activité : Au début de la séance, on fait un petit rappel sur les deux outils présentés dans la séance précédente en proposant un exercice de requête sur le corpus Scientext.

Explication : Cette étape permet de rappeler la présentation de Scientext et de Dicorpus. Dans un deuxième temps, l'exercice proposé permet à l'apprenant de se familiariser avec l'utilisation du corpus.

Consigne : Le mot *résultat* est un nom fréquent dans la conclusion, il est intéressant de voir comment il est utilisé dans les constructions verbales. Choisissez les textes qui vous intéressent et faites une requête permettant de repérer les verbes ayant comme sujet *résultat*.

Question : Quel est le verbe le plus fréquent ayant comme sujet *résultat* ? Repérez quelques constructions que vous allez utiliser dans la conclusion de votre mémoire.

### II. Inscription des étudiants sur Scribplus et découverte de l'outil. (20m)

Activité : Avant la mise en rédaction, l'apprenant sera amené à découvrir le logiciel Scribplus qui va l'aider à écrire quelques paragraphes sur la conclusion.

Scribplus :

A quoi sert une conclusion ?

Dans le mémoire, la conclusion est considérée comme aussi importante que l'introduction. Dans la séance précédente, nous avons vu que la conclusion est le lieu où l'auteur procède au bilan du travail présenté et que l'ensemble de la réflexion se trouve synthétisé. Pour faire ce bilan, on peut procéder en deux étapes :

- faire un rappel de l'étude présentée
- résumer les résultats obtenus

Informations générales sur l'activité :

type : argumentative

thème : rédiger une partie de conclusion pour votre mémoire de recherche

commentaire : Pensez à réutiliser les constructions verbales que vous avez apprises dans la séance précédente. Bon clavier !

Vous pouvez utiliser la ressource Dicorpus pour chercher les expressions et les exemples.

Étape 1. Vous allez tout d'abord revoir la partie d'introduction rédigée, et puis faire un rappel de l'étude en reformulant le sujet de recherche, la problématique ou la méthodologie adoptée. Développez-les en quatre lignes et n'oubliez pas d'utiliser les constructions étudiées.

Exemple : *Nous avons étudié le jeu stratégique de marché avec bien-monnaie dans le modèle à générations imbriquées. La monnaie est considérée comme un bien durable sans utilité. De ce fait, c'est un modèle proche de celui de Samuelson et il permet une analyse claire des relations entre l'équilibre de concurrence imparfaite et celui de concurrence parfaite. [eco-art-39-conclusion]*

Étape 2. Dans un deuxième temps, vous pouvez faire un retour sur l'hypothèse que vous avez formulée en amont de l'étude. Est-ce que les analyses vous permettent de la valider ? Développez-les en deux ou trois lignes.

Exemple : *Les estimations de notre modèle semblent valider l'interprétation de la relation entre taux du crédit et taux de marché en terme de « facturation du coût des ressources bancaires ». Elles justifient, en particulier, l'hypothèse d'existence d'une structure de financement et de refinancement des crédits nouveaux. <...>Par ailleurs, nos résultats d'estimation semblent conforter l'hypothèse de concurrence imparfaite sur le marché du crédit. [eco-art-443-conclusion]*

Étape 3. Il est temps de mettre en évidence les résultats obtenus en vous demandant ce que vous pouvez déduire de vos analyses. Développez-les en 2 ou 3 lignes.

Exemple : *S'il est permis de valider nos hypothèses à partir du moment où aucun stagiaire en FOAD n'a, à ce jour, abandonné la formation, on peut conclure que la personnalisation de la formation alliée au sentiment d'appartenance à un groupe ont permis de lutter contre les abandons. [sed-com-86-conclusion]*

### **III. Rédaction autonome (60m)**

### **IV. Évaluation et questionnaire (20m)**