



**HAL**  
open science

# Des propriétés syntaxiques des collocations dans des écrits scientifiques

Wioletta Daniel

► **To cite this version:**

Wioletta Daniel. Des propriétés syntaxiques des collocations dans des écrits scientifiques. Sciences de l'Homme et Société. 2016. dumas-01368241

**HAL Id: dumas-01368241**

**<https://dumas.ccsd.cnrs.fr/dumas-01368241>**

Submitted on 19 Sep 2016

**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.



# **Des propriétés syntaxiques des collocations dans des écrits scientifiques**

**DANIEL**

**Wioletta**

Sous la direction de Mme Agnès TUTIN

Laboratoire : Lidilem (Linguistique et Didactique des Langues Etrangères et Maternelles)

UFR LLASIC  
Département Sciences du langage

---

Mémoire de master 2 recherche - 30 crédits - Mention

Spécialité ou Parcours : Linguistique

Année universitaire 2015-2016





# **Des propriétés syntaxiques des collocations dans des écrits scientifiques**

**DANIEL**

**Wioletta**

Sous la direction de Mme Agnès TUTIN

Laboratoire : Lidilem (Linguistique et Didactique des Langues Etrangères et Maternelles)

UFR LLASIC  
Département Sciences du langage

---

Mémoire de master 2 recherche - 30 crédits - Mention

Spécialité ou Parcours : Linguistique

Année universitaire 2015-2016



## **Remerciements**

Je tiens tout d'abord à remercier infiniment la directrice de mon mémoire, Madame Agnès Tutin, de m'avoir fait confiance, de m'avoir guidée, encouragée, conseillée tout au long de la rédaction de ce mémoire. C'était une expérience très enrichissante et je suis reconnaissante d'avoir pu apprendre à ses côtés.

Je remercie également Monsieur Olivier Kraif pour son aide dans les aspects informatiques et l'amabilité de faire partie du jury pendant ma soutenance.

Mes remerciements vont également à Madame Alicja Hajok de l'Université Pédagogique de Cracovie.

Enfin, j'adresse mes plus sincères remerciements à ma famille et à mes amis.

Je souhaiterais remercier mes amies polonaises: Danuta Paryła et Anna Tapek qui m'ont soutenue lors de mon séjour en France, apportant dans les moments de doute de précieuses paroles réconfortantes.

La réalisation de ce mémoire a été possible avec l'aide de la région Rhône Alpes (Bourse CMIRA attribuée pour l'année 2015/2016).

**AUVERGNE – RhôneAlpes\***

### DÉCLARATION

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

NOM : DANIEL ..... PRENOM : KIOLETTA .....

DATE : 05 07 2016 .....



# Sommaire

Introduction .....	6
Partie 1 - La notion de collocation sous différentes perspectives .....	9
<b>CHAPITRE 1. INTERET DU PHENOMENE DES COLLOCATIONS.....</b>	<b>10</b>
1. ETUDES THEORIQUES SUR LES COLLOCATIONS : DE BALLY A NOS JOURS .....	12
Partie 2 - Méthodologie et problématique .....	53
<b>CHAPITRE 2. PROBLEMATIQUE : PROPRIETES SYNTAXIQUES DE COLLOCATIONS.....</b>	<b>54</b>
2. CORPUS ET OUTIL D'EXTRACTION .....	55
3. LE CORPUS SCIENTEXT DES SCIENCES HUMAINS .....	56
4. L'OUTIL DE TRAITEMENT DE CORPUS UTILISE : LEXICOSCOPE.....	57
5. 5 D'ANALYSE PROPOSEE.....	59
Partie 3 - Analyse et proposition de la modélisation .....	69
<b>CHAPITRE 3. ANALYSE DES PROPRIETES SYNTAXIQUES DES COLLOCATIONS ET PISTES DE CODAGE ....</b>	<b>70</b>
<b>CHAPITRE 4. PROPOSITION DE LA MODELISATION .....</b>	<b>80</b>
Conclusion.....	88



## Introduction

Le présent mémoire s'inscrit dans le cadre du projet ANR Termith (Terminologie et Indexation de Textes en sciences Humaines) avec les travaux du laboratoire LIDILEM (Laboratoire de Linguistique et Didactique des Langues Etrangères et Maternelles) de l'Université Grenoble Alpes (Stendhal 3 de Grenoble) sur un lexique de base transdisciplinaire des écrits scientifiques.

Ce mémoire a pour objectif d'observer les constructions syntaxiques des collocations afin de comprendre leur fonctionnement et aussi de proposer leur codage. De cette façon, nous réfléchissons au fait de la représentation et la modélisation des collocations dans des écrits scientifiques. Le but de la modélisation de ces propriétés est de créer un dictionnaire électronique utilisable en ligne. Cette étude relève de la syntaxe et de la linguistique de corpus. Nos analyses porteront exclusivement sur le codage linguistique des propriétés pertinentes à associer à ces collocations (Tutin, 2004) (alternances syntaxiques, détermination, information d'usage) à partir des observations en corpus. Nous sélectionnerons les collocations et nous proposerons une adaptation pour ces expressions. Nous pensons que c'est un sujet intéressant parce que les collocations subissent un éventail de transformations grammaticales. Deux collocations du même type qui, au premier abord, semblent avoir les mêmes restrictions ne subissent pourtant pas les mêmes changements. Nous verrons les exemples dans ce qui suit. Les collocations sont des structures syntaxiques utilisées quotidiennement dans chaque registre de la langue. Ce sujet n'est pas uniquement intéressant du point de vue linguistique, mais également du point de vue didactique. En effet, notre analyse permettra, à notre avis, de proposer un modèle d'un dictionnaire portant sur des propriétés syntaxiques des collocations du type *N1 prép. N2*. Nous essaierons de présenter une modélisation accessible de ce dictionnaire à chaque utilisateur.

D'abord, nous essayerons de définir ce que le terme *collocation* signifie. Les collocations, qu'on définira comme des associations lexicales binaires privilégiées (Cf Tutin 2010 ; Grossmann & Tutin, soumis) ne sont pas uniquement des associations de deux mots sémantiquement pleins.

Il faut également ajouter que les collocations présentent des spécificités au plan syntaxique, qu'il s'agisse de la distribution, des alternances syntaxiques, de la sous-catégorisation (cf. Tutin, A. (2004), Tutin & Kraif (en préparation)). Nous pouvons même

nous demander si certaines de ces propriétés syntaxiques ne constituent pas des critères essentiels à prendre en compte dans la définition même du phénomène.

Dans ce mémoire, nous nous attacherons à explorer ces problématiques, à l'aide des méthodes de la linguistique de corpus. Le champ lexical sera celui des collocations du lexique scientifique transdisciplinaire, pour lequel des extractions ont déjà été effectuées. L'outil d'extraction des collocations et des propriétés sera le Lexicoscope (Kraif *et al.* 2014). Nous réfléchirons ensuite à la modélisation et la représentation lexicographique de ces collocations.

Notre motivation pour les propriétés syntaxiques des collocations est liée à la nouveauté de cette approche. À l'heure actuelle le phénomène des *collocations* pose des problèmes intéressants pour la théorie linguistique. Elles ne sont ni complètement libres ni complètement figées. Dans une perspective didactique, il est indispensable que les dictionnaires fournissent une information linguistique détaillée sur les collocations. Cela peut être spécialement utile pour l'enseignement aux étudiants locuteurs non natifs dans le cadre de la didactique du français de spécialité. L'information syntaxique devrait préciser le type de constructions qui accompagnent le mot (type de complément, etc.), le type de déterminant (qui est souvent complexe à prévoir) et les aspects distributionnels comme la place des adjectifs. Même si les collocations sont plutôt faciles à décoder, il est difficile de les utiliser, ainsi que de prédire leur usage. Nous soutenons qu'il faut combler un vide et mettre en place un bon dictionnaire accessible en ligne. L'objectif de ce mémoire est de fournir des informations nécessaires pour pouvoir créer un dictionnaire électronique utilisable en ligne.

Notre travail se présente sous la forme de trois chapitres. Le premier chapitre explique l'intérêt du phénomène des collocations et présente la notion de collocation sous différentes perspectives. Après un bref historique de la notion, nous observerons qu'elle n'est pas vraiment nouvelle dans plusieurs traditions, mais qu'elle était plutôt traitée sans importance. Ensuite, nous nous pencherons sur les dictionnaires des collocations dans 3 langues vivantes : *l'anglais, le français et le polonais* et sur leur représentation dans ces dictionnaires. Le deuxième chapitre concerne la méthodologie et la problématique de l'analyse qui est liée à la question des corpus à utiliser. Cette section éclaire la problématique complexe d'extraction automatique des données. Afin d'effectuer une telle analyse, nous avons passé beaucoup de temps à apprendre l'utilisation des requêtes permettant d'obtenir des informations voulues. De nos jours, les recherches en linguistique posent les défis intéressants imposés par les progrès en linguistique de corpus. Le dernier

chapitre présente l'analyse des résultats et une tentative de modélisation des propriétés syntaxiques pour un dictionnaire de collocations.

## **Partie 1**

-

**La notion de collocation sous différentes perspectives**

## Chapitre 1. Intérêt du phénomène des collocations

La question des collocations est centrale notamment dans les disciplines de la linguistique telles que la sémantique et la syntaxe. Le traitement automatique des langues et la didactique s'intéressent également à cette notion. C'est un sujet très intéressant du point de vue d'un étranger ou d'un enseignant FLE (Français Langue Etrangère). Cela peut être spécialement utile en vue des objectifs didactiques : pour l'enseignement aux étudiants locuteurs non natifs dans le cadre de la didactique du français de spécialité. En règle générale, les collocations ne constituent pas un défi important pour la compréhension en langue étrangère ou dans l'analyse automatique. Cependant, ces associations privilégiées sont souvent peu prédictibles, ainsi, difficiles à reproduire. Il faut dire qu'il y a plus de collocations que les mots – fait peu surprenant, car les mots se produisent dans les collocations différentes. L'un des principaux obstacles concernant l'enseignement des collocations concerne leur nombre qui peut s'élever même à des dizaines des milliers. Néanmoins, l'objectif principal n'est pas de les apprendre toutes par coeur. Il faut se focaliser avant tout sur la sélection des mots et des collocations.

La question la plus importante qui se pose dans l'enseignement des collocations est la suivante : comment présenter la forme et le sens des éléments afin d'aider à leur mémorisation et leur utilisation ? Nous avons pu observer une grande importance accordée à l'enseignement de la grammaire. Les apprenants doivent retenir beaucoup de règles grammaticales avant même de comprendre ce qu'elles signifient. Quant à la composante lexicale, l'enseignement/apprentissage des langues étrangères se réduit très souvent à la mémorisation d'une liste de mots isolés et hors contexte. Dans ce qui suit, nous soutenons l'hypothèse qu'avant tout, il faut apprendre des collocations en bloc, parce que seulement le bloc permet de comprendre le sens de l'ensemble (Lewis, 2000) Autrement dit, il est plus facile de retenir des « blocs de mots » que de rassembler des mots isolés.

<b>Mots simples</b>	<b>Collocations</b>
(une) Question	Poser une question
(une) Table	Débarrasser/mettre/... une table

(une) Hypothèse	Faire/avancer/écrire/... une hypothèse
...	...

Tableau 1 : Exemple de <bloc de mots>

Les études menées par Nadja Nesselhauf (2005) ont montré que malgré quelques années d'apprentissage d'une langue étrangère, elles n'ont pas d'effet positif pour l'apprentissage des collocations. C'est un point très important dans la didactique moderne. La question du choix lexical est la plus importante. Le choix antérieur de certains mots entraîne mécaniquement la sélection d'autres mots (Polguère, 1998). Il faut bien souligner que certaines associations plus idiomatiques sont préférées. Néanmoins, dans l'enseignement, le phénomène de collocations est encore négligé, même ignoré. Avant tout, il faut d'abord sensibiliser les apprenants à la notion de *collocation*. Il est important que l'apprenant sache qu'il y a des associations privilégiées dans la langue qui se combinent entre elles. Dans ce cas-là, les dictionnaires devraient venir à l'aide, mais nous allons montrer que parfois ce n'est pas évident. Nous allons faire un inventaire des dictionnaires des collocations dans ce qui suit. L'apprenant devrait apprendre comment utiliser ces dictionnaires afin d'acquérir des connaissances à utiliser les collocations (Nesselhauf, 2005). D'abord, il faut savoir comment enseigner l'apprentissage des collocations. Il y a une hypothèse formulée qui dit que l'enseignement explicite des collocations et la pratique d'usage favorise l'emploi de ces expressions. Dans son livre *Collocations in a Learner Corpus*, Nadja Nesselhauf a comparé deux pays et deux modes d'apprentissage : celui de Pologne et celui de l'Allemagne. Elle a démontré que les étudiants polonais utilisent plus souvent les expressions adaptées par rapport aux étudiants germaniques. Elle croit que la différence se trouve dans l'apprentissage. En Allemagne, ils préfèrent l'approche communicative qui est basée sur la créativité alors qu'en Pologne, ils se focalisent sur la précision dans la langue. Chaque approche a ses bons et ses mauvais côtés, nous n'allons pas maintenant en discuter. Nesselhauf propose de trouver un juste milieu entre ces deux approches. Pour les apprenants avancés, elle émet une hypothèse que les critères de la sélection des collocations peuvent se baser sur la fréquence et le degré de difficulté. Elle souligne également que l'apprentissage doit être systématique.

Finalement, ce qu'il faut encore souligner, c'est le fait qu'il faut bien distinguer dans l'apprentissage de FLE *le vocabulaire actif* qui sert à la production des textes et *le vocabulaire passif* qui permet de comprendre les textes au moment de leur réception. **Les collocations sont le vocabulaire à apprendre activement.** (Hausmann, 2002)

## 1. *Etudes théoriques sur les collocations : de Bally à nos jours*

Pour notre travail sur les propriétés syntaxiques des collocations dans des écrits scientifiques, nous allons d'abord aborder la question de la phraséologie et la place des collocations dans les modèles linguistiques.

Même si la notion de *collocation* – généralement utilisée sous d'autres appellations – n'était pas nouvelle pour les lexicographes ou les didacticiens de la langue, elle n'était pas vraiment bien traitée pendant les années. En d'autres termes, c'était une notion ancienne, mais vraiment peu théorisée. Depuis quand la phraséologie a commencé à occuper une place importante dans la linguistique, nous avons observé un nouvel essor et un élargissement des objets. Traditionnellement, la phraséologie relève d'une approche syntaxico-sémantique et combinatoire du lexique. Les nouvelles recherches dans le domaine portent sur le « principe phraséologique du langage » (Legallois, Tutin, 2013) et intègrent des objets tels que les collocations, les schémas syntaxiques ou textuels. Autrement dit, la phraséologie a étendu les objets d'études. Auparavant, elle s'occupait plutôt des locutions phraséologiques.

On retrouve des traces des mentions de la notion de la collocation dans la tradition française à partir du XVIIe siècle. Nous n'allons pas refaire un long historique, vu que la dernière décennie a apporté des approches innovatrices sur la question des associations lexicales, en particulier, des collocations, ainsi, nous allons nous pencher sur des tendances nouvelles.

Les collocations sont des mots qui tendent à apparaître ensemble. Pour mieux illustrer la problématique, nous pouvons apercevoir que l'intensité dans *remercier* se traduit de manière préférentielle par *infiniment* (*remercier infiniment*), pendant que le verbe causatif standard qui est normalement associé à *peur* est *faire* (*faire peur*) (Tutin, 2004)

La notion de phraséologie est répandue dans le monde et elle est définie différemment selon le concepteur (Meunier, Granger, 2008). La linguistique du corpus a contribué également à l'extension du champ phraséologique. Nous allons nous axer sur la notion de *phraséologie étendue*. Comme nous l'avons dit ci-dessus, la phraséologie intègre des objets d'étude variés, allant des collocations aux séquences discursives (Legallois, Tutin, 2013). Les objets de la phraséologie étaient auparavant perçus comme des anomalies ou des exceptions. Aujourd'hui, ils deviennent des éléments centraux dans les modèles linguistiques où la notion de principe phraséologique du langage se développe.

Les collocations se trouvent à la frontière de la combinatoire libre et de la combinatoire figée. En linguistique, il y a deux approches concernant les collocations : l'approche *quantitative* et *qualitatives* (Hausmann, Blumenthal, 2006). Les études quantitatives se basent sur la description statistique de corpus. Il est question de retenir des cooccurrences statistiquement fréquentes. Quant aux études qualitatives, il se veut que les collocations se définissent comme une conception « restreinte » de la collocation, entendue comme une association lexicale syntagmatique restreinte entre deux éléments ayant une relation syntaxique. C'est avant tout la définition qui est issue de l'approche continentale et qui nous semble la plus adéquate pour les objectifs de ce travail. Il faut bien souligner que les deux approches ne s'opposent pas.

Comme nous l'avons mentionné ci-dessus, un des problèmes assez importants rencontrés aujourd'hui est que l'on trouve des définitions très variées pour désigner les mêmes phénomènes. Pour mieux illustrer la problématique, il faut savoir que certaines choses reçoivent d'autres appellations selon le linguiste. Charles Bally (1865 – 1947, un linguiste suisse), utilise le terme de « séries phraséologique » pour parler de *collocation* tandis qu'Igor Mel'čuk (1932, un professeur et linguiste) parle de « semi-phrasème ». Il serait donc indispensable de choisir au départ un cadre théorique et terminologique qui permettra de définir avec rigueur les principaux types de phraséologismes. Pour les objectifs de ce travail, nous essayerons de définir ce que le terme *collocation* signifie.

1865 - 1947	Charles BALLY	les séries phraséologique (d'après Bally, 1951) = <i>collocations</i> qu'il appelle aussi « groupements usuels »
1921 - 2002	Eugen COSERIU	L'inventaire du « discours répété » où les « analyses préférées » ( <i>collocations</i> ) apparaissent parmi les combinaisons figées de mots
2003	Geoffrey WILLIAMS	L'école britannique - le rôle du contexte du terme de <i>collocation</i>
	Robert GALISSON	L'analyse des <i>collocations</i> en un terme-noyau [= la base] et un terme-satellite [=collocatif]
1943	Franz Joseph HAUSMANN	L'inspiration de Bally et de Galisson



1932	Igor MEL'CUK	Fonctions lexicales + analyse collocationnelle
	Francis GROSSMAN, Agnès TUTIN, Alain POLGUERE	L'essai de proposer une définition (plus) claire de la notion

Tableau 2 : un tableau récapitulatif

Comme nous l'avons dit ci-dessus, la linguistique du XIXe siècle ignore les collocations. C'est d'abord Charles Bally, disciple de Saussure qui dénomme pour la première fois cette notion. Il distingue deux types de « locutions phraséologiques » [= phrasèmes] : les « unités phraséologiques », appelées autrement les locutions idiomatiques et « les séries phraséologiques » = [collocations] qu'il appelle aussi les « groupements usuels » et pour lesquels il cite de nombreux exemples : *grièvement blessé, gravement malade, désirer ardemment*, etc. (Hausmann, Blumenthal, 2006 : 7, d'après Bally, 1951). Selon Charles Bally, dans ces « séries phraséologiques », les éléments conservent leur autonomie, mais ils présentent aussi des « affinités » pour lesquelles les prototypes sont les séries d'intensité (*chaleur suffocante*) ou les périphrases verbales (*prendre une décision*). Nous apercevons que les adverbes ne sont pas facilement interchangeables dans les séries phraséologiques. Nous ne pouvons pas dire *\*grièvement malade* ou *\*désirer éperdument (aimer éperdument)*.

Entre les extrêmes (groupements passagers et unités indecomposables) se placent des groupes intermédiaires appelés séries phraséologiques (p.ex. les séries d'intensité ou les périphrases verbales). (1909 : 66)

Il y a série ou groupement usuel lorsque les éléments du groupe conservent leur autonomie, tout en laissant voir une affinité évidente qui les rapproche, de sorte que l'ensemble présente des contours arrêtés et donne l'impression du « déjà vu ». (1909 : 70)

Comme nous l'avons déjà dit dans la partie précédente, Charles Bally distingue bien deux types de « locutions phraséologiques », ainsi, il distingue des collocations des locutions.

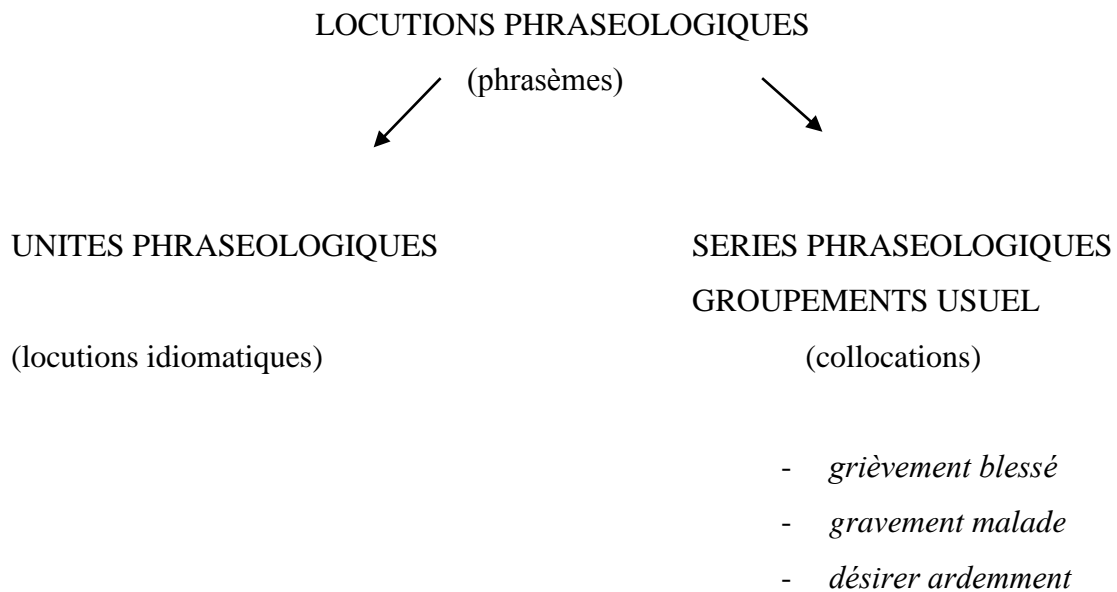


Tableau 3 : schéma de Bally

Ensuite, un grand linguiste de philologie romane, Eugenio Coseriu parle de « clichés lexicaux » et « syntagmes lexicaux » (Hausmann, Blumenthal, 2006 : 8) en prenant comme des exemples *un gros chagrin* et *une grande douleur*.

Avec l'école britannique et le terme introduit par Geoffrey Williams, nous découvrons le rôle du contexte du terme de collocation (Williams, 2003). Sa conception unit les collocations à des associations lexicales fréquentes où les éléments ont des relations syntaxiques. C'est une conception plus étroite qui, parfois, est appelée *collocation lexicale restreinte*.

Robert Galisson, un linguiste et un professeur émérite, s'inspire des travaux britanniques et il analyse les collocations en *un terme-noyau (=la base)* et *un terme-satellite* qu'il appelle *collocatif* (Hausmann, Blumenthal, 2006 : 8, d'après Galisson 1971 [1965 : 14]).

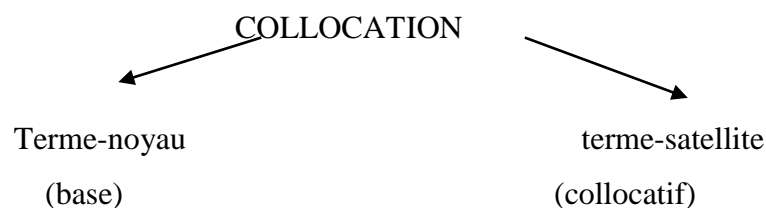


Tableau 4 : schéma de Galisson

Franz Joseph Hausmann, un linguiste, s'inspire de Ch. Bally et de R. Galisson. Il perçoit également la *collocation* comme **une construction constituée d'une base et d'un collocatif**.

Quant à lui, il définit la *collocation* comme une combinaison phraséologique, codée en langue qui se compose d'une base et d'un collocatif. Suivant son raisonnement, nous voyons que la base est un mot (c'est-à-dire, l'acception d'un mot, appelée aussi « lexie ») que le locuteur choisit librement parce qu'il est souvent traduisible et apprenable sans la base. Prenons comme exemple : *grièvement blessé* et *passer un examen*. La base (dans ce cas-là : *examen*, *blessé*) et le collocatif (ici : *passer*, *grièvement*). Le collocatif est un mot (ou, autrement dit, l'acception d'un mot) que « le locuteur choisit en fonction de la base parce qu'il n'est pas définissable, traduisible ou apprenable sans la base. » (Blumenthal, Hausmann, 2006 : 4).

Hausmann fait une distinction entre la combinaison libre et la combinaison phraséologique. Pour lui, les relations collocationnelles concernent une « relation binaire polaire » dans le sens où un argument dans la relation domine sur l'autre. Hausmann a parlé de la classe lexicale ouverte (verbes, noms, adverbes et adjectifs). Il a suggéré de classer les relations collocationnelles selon la partie du discours des lexèmes entre lesquelles ces relations sont entretenues (Wanner, 1996). Il a fait une classification des collocations :

Table 1: *Classification of collocations following Hausmann*

Collocation class	Examples
V + N	[to] <i>have a chat</i> [to] <i>receive attention</i>
V + ADV	[to] <i>play fair</i> [to] <i>complain bitterly</i>
N + N	<i>dance marathon</i> <i>marriage ceremony</i>
A + N	<i>deep regret</i> <i>beguiling smile</i>
ADV + A	<i>highly competent</i> <i>deep black</i>

Tableau 5 : schéma de Hausmann

Néanmoins, cette classification n'est pas bien détaillée parce qu'il existe encore d'autres types de collocations.

En règle générale, un locuteur étranger perçoit l'apprentissage des bases comme l'apprentissage d'équations simples. Néanmoins ce raisonnement est faux. Prenons quelques exemples dans 3 langues européennes : ang. *exam*, pol. *egzamin*, franç. *examen*. Nous savons que la valeur du collocatif varie en fonction de la base (pour dire *passer en examen*, en anglais, on dit *take < prendre >*, en polonais *zdawać < passer, réussir >* et le collocatif anglais *pass* veut dire en français *< réussir >*, le collocatif polonais signifie *< passer >*.

FR : EXAMEN

ANG : EXAM

PL : EGZAMIN

FR : passer un examen

ANG : take an exam

PL : zdawać egzamin

Si un locuteur fait un calque de français en parlant des examens et au lieu de dire *j'ai passé un examen*, il dit *I passed the exam*, le message est confondu. *Passer* dans ce cas-là ne signifie pas *réussir* mais *take*. Cela va de même pour la langue polonaise. *Zdawać* c'est *passer* en français et *take* en anglais et *zdać* c'est *réussir* (fr) et *pass* (ang). C'est pour cela qu'il est important d'apprendre la base et le collocatif ensemble.

Il arrive aussi que même un apprenant avancé puisse avoir des difficultés dans une situation de détresse d'encodage où même s'il connaît la base, il cherche le collocatif. Quant à Igor Mel'čuk, il a fait une typologie des phrasèmes : *phrasème complet, quasi-phrasème, semi-phrasème* où il a classé les collocations parmi les *semi-phrasèmes* (Mel'čuk, 1995). Cependant, en 2011, il a présenté une nouvelle typologie. Selon lui, la façon la plus générale de caractériser la notion de phrasème est de dire qu'un phrasème est un énoncé multilexémique non libre. Il explique qu'« un énoncé multilexémique est une configuration de deux ou plus lexèmes syntaxiquement liés ». Dans la phrase : « La façon la plus générale de caractériser la notion de phrasème est de dire qu'un phrasème est un énoncé multilexémique non libre. » (Mel'čuk, 2011 : 1), il y a plusieurs énoncés multilexémiques : *la façon, la plus générale, la façon la plus générale, la façon de caractériser la notion de phrasème*, etc., ainsi que la phrase entière. (Mel'čuk, 2011). Suivant son raisonnement, nous pouvons ajouter qu'« un énoncé multilexémique est libre

si et seulement si [= ssi] il n'est pas contraint sur l'axe paradigmatique ». Autrement dit, il est question de voir si le sens de chacune de ses composantes lexicales est sélectionné par le locuteur, indépendamment des autres composantes. Ainsi, nous voyons que chaque composante d'un énoncé libre (sauf les éléments grammaticaux) peut être remplacée par une expression assez synonyme qui garde le contenu linguistique et le sens de cet énoncé. Cependant, il existe des énoncés qui ne sont pas libres – prenons comme exemple *prendre une décision* (*Jean a pris une décision*) : si le lexème DÉCISION est sélectionné librement, vu que locuteur veut parler de 'décision', le lexème PRENDRE ne l'est pas. Dans cette expression, *prendre* ne peut pas être remplacé par d'autres verbes comme *accepter*, *ramasser*, *saisir*, *toucher*, etc. sans influencer la correction de l'expression (Mel'čuk, 2011). Notons qu'en anglais, on dirait *make* [= 'faire'] *a decision*, en polonais : **podjąć** *decyzję* [= '**prendre**']. Les preuves apportées par ces exemples montrent que PRENDRE est sélectionné en fonction de DÉCISION. Imaginons qu'au lieu de DÉCISION le locuteur utilise un autre mot, par exemple, le CHOIX (*Jean a pris la décision de rester*  $\cong$  *Jean a... le choix de rester*), il aurait dû dire FAIRE plutôt que PRENDRE : *Jean a fait \*pris le choix de rester*.

D'après I.Mel'čuk, cette définition couvre tous les types de phrasèmes et rien que des phrasèmes. Avec la notion proposée par lui, nous apercevons qu'un phrasème ne peut pas être librement construit par le locuteur, ce qu'il comprend par là est que le phrasème doit donc être stocké dans la mémoire. Enfin, il a proposé une typologie de phrasème : *locutions, collocations et clichés*.

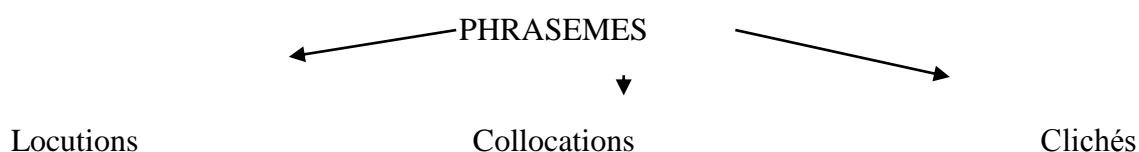


Tableau 6 : schéma de Mel'čuk

### ***1.1. L'école britannique : John Rupert Firth, John Sinclair***

La définition de *collocation* diffère selon le linguiste. Dans cette section, nous allons présenter un autre point de vue qui tire son origine dans l'école britannique. Même si, dans notre mémoire, nous allons choisir la conception de J. Hausmann et I. Mel'čuk, nous trouvons important de présenter d'autres approches concernant cette notion.

J.R Firth - un lexicographe et linguiste qui a instauré la valeur des études fondées sur le Corpus (Clear, 1992) a théorisé la notion de *collocation* dans une perspective textuelle. Il a dit : « You shall know a word by the company it keeps ». Selon lui, (Firth, 1957 : 179) le sens des mots se déduit par l'environnement lexical dans un contexte donné. Les combinaisons possibles mises en jeu par les collocations peuvent être soit des expressions figées (idioms) soit des constructions. Si nous suivons son raisonnement et nous trouvons que le sens des mots se déduit en quelque sorte de leur environnement lexical dans un contexte donné, (« you shall know a word by the company it keeps »), il n'est pas dans les objectifs du linguiste de décrire le sens en dehors du texte, mais seulement dans un registre donné, dans un « restricted text ». Nous trouvons que sa définition n'est pas très claire. De plus, il rejette le fait que la collocation puisse être restreinte aux associations de mots, comme il était dit dans la tradition lexicographique.

Parlons maintenant de John Sinclair (1933 – 2007) – un linguiste britannique qui s'inscrit également dans l'approche britannique. Une différence entre J. Sinclair et J.R Firth est selon le second, la collocation n'est pas limitée par l'association de mots. John Sinclair (1992) dans son livre intitulé *Corpus, Concordance, Collocation* montre une approche empirique à la langue. Il insiste avant tout sur les critères de type statistique :

Collocation is the occurrence of two or more words within a space of each other in a text... Collocations [...] can be important in the lexical structure of the language because of being frequently repeated. This second kind of collocation, often related to measure of statistical significance, is one that is usually meant in linguistic discussions (J.Sinclair, 1991 : 170).

Autrement dit, selon lui, la *collocation* est une association habituelle. Ce sont des mots qui cooccurrent dans un certain ordre et qui vont toujours ensemble, c'est-à-dire, des mots qui tendent à souvent apparaître ensemble et qui sont indissociables de la notion

de corpus. J.Sinclair disait que c'est la récurrence des associations lexicales dans les textes qui leur attribue le statut de collocation (Sinclair, 1991).

John Sinclair s'intéresse avant tout à la linguistique de corpus. Il développe dans son livre *Corpus, Concordance, Collocation* (1991) l'idée de l'utilisation des corpus pour l'étude de la linguistique. Dans l'ensemble, il oppose la notion de *colligation* à la notion de *collocation*. Selon lui, la *colligation* renvoie à l'environnement grammatical privilégié du mot tandis que la *collocation* (Legallois, 2012), comme nous l'avons déjà mentionné ci-dessus, est une expression récurrente.

Nous pouvons alors remarquer que la linguistique fonctionnelle anglaise, avec les travaux, entre autres de J. Sinclair, a accordé une place importante aux questions phraséologiques. John Sinclair a proposé d'associer les niveaux *syntaxique* et *lexical*, autrement dit, il ne voulait pas séparer ces deux niveaux. Il trouvait que de nombreux mots ne sont pas choisis librement, mais sont déjà présents dans des structures préfabriquées. Ainsi, nous ne pouvons pas dissocier le lexique et la syntaxe parce que ce serait artificiel. Ces deux niveaux s'entremêlent.

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 be analysable into segments (J.Sinclair, 1991 : 110).

Nous voyons bien que la notion de *collocation* qui sort de la perspective fonctionnelle est centrale. Il définit la *collocation* comme des associations de mots (y compris les mots grammaticaux) qui apparaissent souvent en cooccurrence, mais qui n'ont pas nécessairement de relation syntaxique directe.

Collocation is the cooccurrence of two or more words within a short space of each other in a text. The usual measure of proximity is a maximum of four words intervening. (Sinclair 1991 : 170).

Son approche est différente à celle proposée par Hausmann et Mel'čuk, qui opposent dans les collocations la base (*remercier* dans *remercier infiniment*) au collocatif (*infiniment* dans *remercier infiniment*). Selon lui, aucun élément dominant n'apparaît dans cette structure. De notre côté, nous suivons la définition proposée par I. Mel'čuk qui définit la *collocation* comme *une cooccurrence lexicale privilégiée de deux éléments linguistiques entretenant une relation syntaxique*.

Suivant son raisonnement, nous essayerons de définir la notion de *collocation*. Étant donné que beaucoup de choses ont été déjà dites dans les travaux précédents, nous allons nous pencher avant tout sur les critères du linguiste russe Igor Mel'čuk.

- **La transparence et le non – figement sémantique de la collocation** : Selon I. Mel'čuk, dans certains cas de figure, le sens n'est pas transparent. Le fameux exemple cité ici concerne l'exemple *peur bleue*. L'adjectif *bleue* dans ce cas-là est un intensifieur qui renforce *peur* et ainsi la collocation signifie *avoir très peur*. Néanmoins, rien dans l'adjectif *bleu* ne nous permet de supposer s'il s'agit d'une grande peur ou d'une peur immense.
- **Le caractère binaire de la collocation** : en règle générale, la collocation se compose de deux mots, prenons comme l'exemple *fort comme un turc* : nous voyons ici l'association d'un mot (*fort*) avec un syntagme (*comme un turc*). Néanmoins, il faut ajouter que la question de la binarité n'est pas si simple. Les définitions du terme de collocation dans la tradition continentale la décrivent comme une expression binaire (Mel'čuk, 1998).

A COLLOCATION AB of language L is a semantic phraseme of L such that its signified « X » is constructed out of the signified of one of its two constituent lexemes – say, of A – and a signified « C » [« x » = 'AoC'] such that lexeme B expresses only 'C' contingent on A. (Mel'čuk, 1998).

A. Tutin reformule de façon plus précise le critère de binarité catégorielle comme « La collocation met en jeu une unité lexicale simple ou complexe en relation syntaxique avec un constituant linguistique » (Tutin, 2013). Suivant son raisonnement, nous voyons les définitions qui précisent que les deux éléments d'une collocation n'ont pas le même statut : la base conserve en général son sens habituel pendant que le collocatif est dépendant de la base.

- **La notion de cooccurrence restreinte** : chez I. Mel'čuk la *collocation* est plutôt envisagée dans le cadre de la production. Selon lui, le choix n'est pas libre, mais toujours imposé, ce que nous pouvons entendre par là est que la cooccurrence est restreinte. Pour éclaircir ce concept revenons à l'exemple cité ci-dessus : pour intensifier *peur*, le locuteur choisira préférentiellement le mot *bleue*.



Voici la définition proposée par Igor Mel'čuk :

A COLLOCATION **AB** of T is a semantic phraseme of T such that its signified 'X' is constructed out of the signified of the one of its two constituent lexemes - say, of **A** - and a signified 'C' ['X' = 'A  $\wedge$  B'] such that the lexeme **B** expresses 'C' contingent on **A**.

The formulation " **B** expresses 'C' contingent on **A** " covers four major cases, which correspond to the following four major types of collocations :

1. **either** 'C' = 'B', i.e. **B** does not have (in the dictionary) the corresponded signified;  
**and** [ a. 'C' is empty, that is, the lexeme **B** is, so to speak, a semi-auxiliary selected by **A** to support it in a particular syntactic configuration;  
**or b.** 'C' is not empty but the lexeme **B** expresses 'C' only in combination with **A** (or with a few other similar lexemes)];
2. **or** 'C' = 'B', i.e., **B** has (in the dictionary) the corresponding signified;  
**and** [ a. 'B' cannot be expressed with **A** by any otherwise possible synonym of **B**;  
**or b.** 'B' includes (an important part of) the signified 'A', that is, it is utterly specific, and this **B** is 'bound' by **A**].

**Examples** (lexeme **A** is in small caps)

Case **1a** : collocations with support (= 'light') verbs, such as [to] do a FAVOR, [to] give a LOOK, [to] take a STEP, [to] launch an APPEAL, [to] lay SIEGE [to N].

Case **1b** : collocations such as black COFFEE, French WINDOW, Fr. BIÈRE bien frappée <\*battue> ...

Case **2a** : collocations with intensifiers, such as strong <\*powerful> COFFEE, heavy <\*weighty> SMOKER, deeply <\*profoundly> MOVED, [to] ILLUSTRATE vividly.

Case **2b** : collocations such as The HORSE neighs, aquiline NOSE, rancid BUTTER or artesian WELL.<sup>1</sup>

Essayons d'analyser à notre égard sa définition<sup>2</sup>. Comme il peut être vu au-dessus, la collocation est l'association d'un mot simple ou d'une lexie<sup>3</sup> L et d'un constituant C qui ont entre eux une relation syntaxique :

---

<sup>1</sup> Mel'čuk, I. (1998). Collocations and lexical functions. en Anthony Paul COWIE (ed.)(2001 [1998]), 23-54.

<sup>2</sup> Inspiré par : Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif Agnès Tutin, Francis Grossmann

<sup>3</sup> Selon CNRTL : LING. Unité lexicale de langue constituée soit par un mot (*lexie simple*) soit par des mots associés (*lexies composée et complexe*).

- Le collocatif C exprime un sens donné en cooccurrence avec la base L
- Le sens de L est habituel

Étant donné que notre petite étude n'est qu'une ébauche et que le sujet est vraiment très vaste, nous omettons l'opposition entre la collocation opaque et transparente, de ce fait, nous n'entrons pas non plus dans les détails concernant la collocation dite régulière. Cette typologie ne nous paraît pas essentielle puisque nous allons nous appuyer sur la typologie melc'ukienne. Notre objectif n'est pas d'en discuter parce qu'il existe déjà une bibliographie relativement complète à ce sujet. Dans ce travail, nous nous focalisons sur la collocation de façon générale et nous essayons de choisir un cadre théorique qui nous permettra d'avancer dans le sujet.

### ***1.2. Propriétés syntaxiques des collocations***

La notion de collocation peut être définie de façon différente. Certaines personnes argumentent que ce terme dénote une cooccurrence lexicale privilégiée et présente des particularités syntaxiques qui varient selon le statut des associations. Si nous voulions élargir ce concept, nous dirions que les collocations appelées « régulières » (Tutin, 2004) ont le moins de contraintes sur le plan syntaxique.

Notre intérêt pour les collocations prend racine dans la pratique langagière et l'usage de ces expressions par des apprenants de la langue française. Les collocations sont présentes dans la langue et même des apprenants avancés ont parfois des difficultés à s'en servir. Par ailleurs, il faut avoir conscience que certaines associations sont préférables dans une langue pour ne pas faire de calques de sa langue maternelle. Cela se vérifie aussi dans les langues de spécialité. Les collocations sont fixées par l'usage, fondées sur des connotations et des conventions idiosyncrasiques. Cela paraît difficile pour un apprenant de langue étrangère qui n'a pas de moyen de pressentir l'acceptabilité d'une collocation et ainsi, qui s'appuie naturellement sur les modèles de sa langue maternelle.

Nous concentrons notre attention sur des spécificités au plan syntaxique. Nous nous intéresserons à la distribution, aux alternances syntaxiques et à la sous-catégorisation (cf. Tutin, (2004), Tutin & Kraif (en préparation)).

En règle générale, les linguistes séparent le lexique de la syntaxe. Celle-ci est une branche de la linguistique qui s'occupe de la combinaison des mots au sein d'une phrase. Grâce à elle, nous sommes capables de créer des énoncés grammaticalement bien normés.

La syntaxe est perçue par le linguiste Watters John R (2004) comme une analyse des mots et des morphèmes entrant en relation l'un avec l'autre afin de former des phrases grammaticales. Du point de vue syntaxique, les linguistes étudient un ordre linéaire des mots constituant les unités structurées supérieures (les syntagmes et les propositions) qui forment des structures hiérarchisées (Watters, 2004 : 231).

Autrement dit, la syntaxe représente donc un ensemble de règles dans lesquelles s'insèrent les éléments du lexique. Comprendre le fait qu'il est indispensable pour la grammaire d'une langue de savoir comment se combinent les mots sur le plan syntaxique peut aider à dévoiler certaines restrictions. Notre analyse s'occupe de la constructions des collocations. Elle traite donc la question de la syntaxe étant donné que la collocation est une combinaison des mots qui apparaissent toujours ensemble.

Prenons maintenant en considération la notion de la sous-catégorisation. Elle est constamment étudiée par nombreux linguistes, entre autre, par Richard Carter. D'après lui, la sous-catégorisation se trouve entre la syntaxe et la sémantique. Elle a pour objectif de décrire chaque entrée lexicale en tenant compte de son contexte. Elle permet d'analyser où un mot est placé dans la phrase et quel sens il prend dans le contexte donné, ce qui nous donne la possibilité d'établir ses traits caractéristiques. (Carter, 1984 : 181). La sous-catégorisation est donc importante parce que les arguments sont en général plus contraints sur le plan syntaxique et sémantique vu leurs restrictions sélectionnelles. Les relations de sous-catégorisation peuvent être les suivantes : la relation objet (*confirmer une analyse*), la relation sujet (*les résultats infirment*), les compléments nominaux en de, que le nom soit recteur (ou tête) (*l'analyse des données*) ou régi (*l'efficacité de la méthode*) (Tutin, 2007 : 7). Il y en a bien d'autres aussi.

Ensuite, nous pouvons mentionner des restrictions syntaxiques. Quelques-unes ont été relevées par Tutin & Grossmann (2002) :

- Construction de façon atypique par rapport aux règles standards de la langue, cf. absence de déterminant dans les constructions à verbe support, qui n'est d'ailleurs pas systématique

- Moins de flexibilité dans le comportement syntaxique au niveau de la distribution syntaxique. Exemple : collocatif épithète à la fonction prédicative ne peut pas toujours apparaître comme attribut, alors que dans les combinaisons libres, cela est toujours possible (exemple : peur bleue/ \*peur est bleue et peur immense/peur est immense).
- Alternance syntaxique possible pour les collocations verbales : contrainte de déterminant; l’effacement du V support dans certains contextes, etc.

Quand nous observons le voisinage entre mots, nous pouvons apercevoir que le type de relation entre eux est pourvu d’une influence sur la qualité des associations produites.

*Les collocations permettent plus de modifications. Elles suivent les règles syntaxiques normales et sont plus libres que les idiotismes. Sur le plan de la variation syntaxique, elles sont similaires aux combinaisons libres. (Laurens, 1999).*

En français, nous arrivons à distinguer différents types de collocations. Dans ce qui suit, nous présenterons une classification syntaxique des collocations dans la langue française.

### ***1.3. Structures des collocations***

En s’appuyant sur la structure morphosyntaxique des collocations, nous pouvons arriver à en distinguer différents types. Il y en a évidemment bien d’autres, mais pour les objectifs de ce travail, nous présentons ce modèle raccourci.

<b>TYPE (modèle des collocations)</b>	<b>EXEMPLE</b>
Nom + Adjectif (N+ADJ)	champ scientifique
Nom + préposition + Nom (N+prép+N)	champ d’étude
Préposition + Nom (Prép+N)	dans le cadre de l’analyse
Verbe + Nom (V+N)	faire un constat

Verbe + Adjectif (V+ADJ)	définir précisément
Adjectif + Adverbe (ADJ+ADV)	gravement malade

Tableau 7 : Différents type de collocations

Hausmann spécifie les structures syntaxiques impliquées quand il parle des collocations :

« On appellera collocation la combinaison caractéristique de deux mots dans une des structures suivantes : a) substantif + adjectif (épithète) ; b) substantif + verbe ; c) verbe + substantif (objet) ; d) verbe + adverbe ; e) adjectif + adverbe ; f) substantif + (prép.) + substantif » (Hausmann 1989, 1010).

Une définition de la collocation sur laquelle nous nous appuyons dans ce travail est celle proposée par Igor Mel'čuk. Rappelons que selon lui, la collocation est une expression bipartite dans laquelle l'un de deux constituants est choisi librement pour exprimer le sens global de l'expression, alors que l'autre est choisi en fonction du premier constituant.

#### Types de variations morphologiques et syntaxiques

Il faut bien dire que certains verbes des collocations peuvent subir des variations de personnes/nombres. Il y a aussi des noms qui se mettent facilement au pluriel, c'est le cas par exemple d'*avoir une/des appréhensions*. Sur le plan syntaxique, certaines collocations relèvent une souplesse dans les alternances alors que d'autres expressions le refusent. Il faut mentionner que dans les textes, les collocations peuvent subir de plusieurs variations. En basant sur la typologie proposée dans Segond & Breidt (1995) citée par Tutin (2013), nous pouvons observer les **variations lexicales et synonymiques** de collocatifs qui peuvent parfois changer le registre. À titre d'exemple prenons une collocation *mourir d'ennui*. Dans ce cas-là, le verbe *mourir* peut être remplacé par un autre verbe, par exemple *périr*. Le paradigme est limité, mais ouvert. Cela va de même pour des variantes dans les prépositions introduisant les arguments (*être paralysé de/par la peur*) ou dans l'absence ou la présence de déterminants (*avoir peur, avoir une peur bleue, etc.*). Ce sujet était déjà bien discuté par certains linguistes.

Il y a également des **variations morphologiques**, c'est-à-dire, des variations de personne/nombre.

Par exemple :

- *recueil de données* vs. *recueil des données* – la première expression représente la structure syntaxique *N1 prép.+(dét.) N2* où l'article indéfini « des », qui devrait accompagner le N2, disparaît en raison de la présence de la préposition « de ». La construction renvoie donc aux données de façon plus globale sans les préciser. Par contre, la deuxième suite des mots se réfère aux données qui sont plus concrètes, plus définies, ce qui explique l'utilisation de l'article contracté « des » (de+les) ;
- *ensemble de pratiques* vs. *ensemble des pratiques* – c'est une autre structure *N1 prép.+(dét.) N2* où l'article peut être, soit présent, soit absent devant le N2. Dans le premier exemple *ensemble de pratiques*, le substantif au pluriel « pratiques » est présenté de manière générale, par contre, dans le deuxième, son référent est plus concret.

Nous ne pouvons pas oublier les **variations distributionnelles**. Il faut se rendre compte que la position des collocatifs adjectivaux et/ou adverbiaux n'est pas gelée. Ainsi, ils peuvent apparaître dans plusieurs positions. Dans ce cas-là, nous parlons des variations distributionnelles. Comme nous l'avons déjà vu, certains adjectifs sont assez souples dans leur emploi (*une immense anxiété, une anxiété immense, son anxiété était immense*) et d'autres sont plus restreintes (*\*une bleue peur, une peur bleue, \*sa peur était bleue*).

Quant aux **alternances syntaxiques**, dans les ressources lexicales, les collocations sont souvent représentées à la voix active, mais elles peuvent bien apparaître aussi dans les constructions à la voix passive (*elle était paralysée de/par la peur*), les constructions pronominales (*se ronger d'angoisse, l'inquiétude s'accroît...*), les phrases relatives (*la peur qu'il ressentait*) (Tutin, 2004)

Nous pouvons dire que les collocations subissent d'une certaine souplesse de fonctionnement. Pour mieux illustrer le concept, prenons comme l'exemple une expression *jouer un rôle*. Sa structure syntaxique (de type verbe-objet) n'est pas figée. Néanmoins, il y a des contraintes de nature lexicale qui apparaissent : nous ne pouvons pas faire d'alternance synonymique et changer le mot *rôle* pour *fonction* *\*jouer une fonction*.

Ainsi, nous voyons que les collocations relèvent de la composante lexicale et à la fois de la composante syntaxique.

Ces aspects syntaxiques sont importants et dans les parties suivantes concernant les ressources, nous nous intéresserons particulièrement au codage de ces propriétés syntaxiques.

#### ***1.4. Les collocations dans les dictionnaires***

Auparavant, les dictionnaires bilingues ont occupé une place importante dans des études linguistiques. Après, les dictionnaires monolingues ont dominé cette place (Verlinde, Bonon, Selva, 2006).

Aujourd'hui, beaucoup de linguistes travaillent sur des corpus informatisés qui sont souvent lemmatisés et informatisés. L'utilisation des corpus permet de voir l'environnement d'un mot de base, ainsi que de contrôler empiriquement les résultats d'un travail théorique sur les contraintes combinatoires (Hausmann, Blumenthal, 2006). Il faut dire que les corpus numérisés ont révolutionné la recherche linguistique et ont augmenté les possibilités d'observations. En règle générale, les dictionnaires de collocations fournissent pour chaque mot de base une liste de mots-satellites présélectionnés (Hausmann, Blumenthal, 2006).

Il existe déjà des dictionnaires de collocations, tout d'abord dans la langue anglaise parce que la problématique de *collocation* a été largement traitée par la linguistique anglaise, ensuite pour d'autres langues européennes comme le français ou le polonais.

Dans le cadre de la lexicographie contemporaine, ce sont surtout les besoins concernant l'enseignement et l'apprentissage des langues étrangères qui sont à l'origine de la conception des dictionnaires de collocations.

Notre étude est une fenêtre sur la compréhension de la modélisation des collocations dans des dictionnaires. Comprendre ce fait peut aider à dévoiler les restrictions liées à ce problème, ainsi, à réfléchir à créer un dictionnaire des collocations en ligne.

Pour nous faire une idée plus concrète, nous avons comparé quelques dictionnaires de collocations dans 3 langues vivantes : anglais, français et polonais.

Parmi les dictionnaires anglais, nous pouvons citer entre autres :

- *Oxford Collocations Dictionary for Students of English* (2002) (désormais OCDSE) (Crowther *et al.* 2002, réédité et augmenté en 2009, McIntosh. Oxford : Oxford University Press) – ce dictionnaire répond aux critères quantitatives.

## OXFORD Collocations | dictionary for students of English

### **decision** *noun*

**ADJ.** **big, crucial, fateful, important, key, landmark** (*law*), **major, momentous** *It was a big decision to have to make. In a landmark decision, the court agreed to hear evidence from twenty years earlier.* | **difficult, hard, tough** *the difficult decision of whether to go to university or nurse her sick mother* | **firm** *We need a firm decision by Friday.* | **prompt, snap** *I had to make a snap decision about what to do with the money.* | **hasty, knee-jerk, rash** | **final, irreversible, irrevocable** *On Monday, the board of directors will meet to make their final decision. The decision is irreversible.* | **informed** *I need more facts before I can make an informed decision.* | **arbitrary** | **good, rational, right, sensible, wise** | **bad, poor, unwise, wrong** | **collective, joint, unanimous** *In the end, the decision to scrap the project was unanimous.* | **majority, split** | **court, government, etc.** | **investment, policy, etc.**

**VERB + DECISION** **arrive at, come to, make, reach, take** *Key decisions are always taken by the editor.* | **announce, give (sb)** *The committee will give us their decision tomorrow.* | **abide by** *The decision has been made, and we must all abide by it.* | **affirm, uphold** *The management committee upheld her decision to fire two of her staff.* | **reconsider** | **appeal against, challenge** *plans to challenge this decision in the High Court* | **overrule, overturn, quash, reverse** *Nobody has the authority to overrule his decision.* | **defer**

**DECISION + NOUN** **maker**

**PREP.** ~ **about/on** *a decision on her future* > Special page at MEETING

Figure 1 : Article de *decision* dans l'OCDSE

Dans l'OCDSE, la collocation est mentionnée sous l'entrée de la base avec une liste de collocatifs. Ce dictionnaire est un ouvrage très volumineux qui contient 9000 mots les plus courants de l'anglais. Le nombre de collocations dans ce dictionnaire s'élève à 250 000. La taille de corpus utilisé est basée sur le Oxford English Corpus et comprend au total 2 milliards de mots.

Ce dictionnaire fournit une liste de mots différents dont la catégorie grammaticale est identifiée. Nous pouvons observer des mots qui apparaissent avec notre pivot. Ils sont regroupés selon leur catégorie grammaticale. Les différents types de structures sont introduits spécifiquement par un indicateur de catégorie, par exemple « VERBE + decision



» ou « ADJECTIF + decision ». Ces collocatifs sont ensuite sous-catégorisés du point de vue sémantique.

Le nom (ici : decision [fr. décision]) est marqué en bleu en haut de la page. Après, nous avons un indicateur de catégorie qui regroupe, par exemple, les adjectifs qui vont avec le mot donné et les exemples qui montrent l'usage de ces collocations (cf. It was a **big decision** to have to make : ADJ + N). À part les adjectifs et les verbes, il y a également des prépositions qui sont utilisées avec le mot présenté (cf. a **decision on** her future).

Classement syntaxique	Classement sémantique
Explicite avec des étiquettes	Regroupement synonymique à un niveau

Le regroupement des collocatifs par quasi-synonymes dans l'OCDSE permet d'en faciliter le décodage sémantique. Il intègre des schémas syntaxiques avec des bases verbales, adjectivales et de nombreux types de collocatifs.

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ Lisible</li> <li>✓ Les exemples données</li> <li>✓ Un indicateur de catégorie qui facilite l'usage</li> <li>✓ Le regroupement des collocatifs par quasi-synonymes</li> <li>✓ La taille du corpus impressionnante</li> </ul>	<ul style="list-style-type: none"> <li>➤ Malgré tout, l'information fournie dans le dictionnaire reste sommaire</li> </ul>

- *The BBI dictionary of English word combinations* (désormais BBI) de Morton Benson, Evelyn Benson et Robert Ilson (première édition : 1986, réédité en 1993 puis en 1997 et en 2010), Amsterdam/Philadelphia : John Benjamins Publishing Company). Il contient des collocations lexicales et grammaticales.

3. an acrimonious, bitter, heated, sharp, stormy ~  
 4. a lively, spirited ~ 5. a campaign; national; parliamentary; public ~ 6. (a) formal ~ (informal ~)  
 7. a ~ about, on, over; between; with 8. in, during a ~; under ~ (the matter is still under ~) 9. (misc.) a matter for ~

**debate II** v. 1. to ~ heatedly, hotly 2. (D; intr.) to ~ about, on (to ~ about disarmament) 3. (D; intr.) to ~ with 4. (G) (esp. BE) Parliament ~d disarming 5. (Q) we ~d (with ourselves) how to do it

**debit I** n. (BE) direct ~ (to pay by direct ~)

**debit II** v. ("to charge") 1. (D; tr.) to ~ against, to (to ~ a purchase against smb.'s account; to ~ an amount to smb.'s account) 2. (D; tr.) to ~ with (~ her account with the entire amount)

**debit card** n. 1. to accept; issue a ~ 2. a ~ expires (your ~ is no longer valid; it expired last week) 3. by ~ (can I pay by ~?) 4. (misc.) to put a purchase on one's/smb.'s ~

**debris** n. to clear ~

**debt** n. 1. to contract, incur, run up; owe a ~; to get into, go into ~ 2. to collect, recover a ~ 3. to discharge, pay (off), settle; wipe out; work off; write off a ~ 4. to cancel; repudiate a ~ 5. a bad; outstanding, unsettled; toxic, unrecoverable ~; mounting ~s (to write off a bad ~) 6. a business; foreign; gambling; personal, private ~ 7. the national ~ 8. in ~ for; to (he is in ~ to me for a large sum, he is deeply in ~ to me; I'm in ~ to my colleagues for their excellent advice) 9. out of ~ (to stay out of ~) 10. (misc.) a ~ of honor; a ~ of gratitude; deep(ly) in ~ (I owe a ~ of gratitude to my colleagues for their excellent advice)

**debut** n. 1. to make one's ~ (to make one's ~ in society) 2. a professional; stage ~ 3. a ~ as (to make one's ~ as an actor)

**decadence** n. 1. to fall into ~ 2. moral ~

**decant** v. 1. (D; tr.) ("to pour") to ~ from; into, to 2. (fig.) (BE) the staff were ~ed from the old building to the new one

**decay** n. 1. to fall into; prevent; stop ~ 2. dental, tooth ~ 3. radioactive ~ 4. environmental; inner-

**decide** v. 1. to ~ unanimously 2. (d; intr.) ("to make a decision") to ~ against (to ~ against buying a car) 3. (legal) (d; intr.) to ~ against; for, in favor of ("to find for") (the jury ~d for the plaintiff) 4. (D; intr.) ("to choose") to ~ between (it was difficult to ~ between the two of them) 5. (d; intr.) to ~ on ("to choose in favor of") (we have ~d on a new computer) 6. (E) ("to choose") we ~d to stay home 7. (H) (BE) what finally ~d you to stay home? 8. (L) ("to make a decision") she ~d that the children would stay home 9. (Q) ("to make a decision") we could not ~ whether to go out or stay home 10. (misc.) to ~ for oneself (I can't advise you: you'll have to ~ for yourself)

**decision** n. ["act of deciding"] 1. to arrive at, come to, make, reach, take (BE) a ~ 2. (esp. legal) to affirm, uphold; appeal (AE), appeal against; hand down (AE), render a ~ 3. (esp. legal) to overrule, reverse a ~ 4. to defer; reconsider a ~ 5. a big, momentous; crucial; fateful; landmark (esp. legal); weighty ~ 6. a clear-cut; ethical; fair; favorable; good, sensible, wise; popular; just ~ 7. a collective, joint; unanimous ~ 8. a final; firm; irreversible, irrevocable ~ 9. an agonizing, difficult; arbitrary; hasty, rash, snap ~ 10. a bad, poor, unwise; unfair, unjust; unfavorable; unpopular ~ 11. a court; majority; split ("divided") ~ 12. a ~ about, on 13. a ~ to + inf. (the government made the ~ to cut taxes) 14. the ~ that + clause (the government made the ~ that it would cut taxes) 15. (misc.) early ~ (AE; used in the process of admitting students to universities) ["decisiveness"] ["firmness"] 16. to lack ~ 17. of ~ (a man, woman of ~)

**deck** n. ["pack of playing cards"] (AE) 1. to cut; shuffle a ~ 2. to stack a ~ ("to arrange cards dishonestly") (also fig.) ["floor as on a ship"] 3. to swab a ~ 4. an aft, after; flight; lower; main; poop; promenade ~; sundeck; upper ~ 5. below; on ~ (she was on the lower ~; I went up on ~; is standing allowed on the upper ~ of the bus?) 6. (misc.) (usu. fig.) clear the ~s for action! ["device for tape recordings"] 7. a tape ~

Figure 2 : Article de *decision* dans le BBI

Il contient des collocations lexicales et des collocations grammaticales qui indiquent les prépositions qui accompagnent les mots. Dans ce dictionnaire la nomenclature est vaste, il y a 18000 entrées, néanmoins, le nombre de collocations s'élève seulement à 90 000. Ce dictionnaire propose aussi des prépositions comme collocatifs pour les bases nominales : des prépositions pleines par exemple *on decision*, ou des prépositions régies comme *decision of*. Il faut bien souligner que les auteurs du BBI considèrent toutes ces structures comme des collocations grammaticales. Le BBI intègre les prépositions pleines dans les collocations grammaticales. Par ailleurs, du point de vue sémantique, seules les prépositions pleines peuvent avoir un statut de collocatif à part entière, les prépositions régies devant être intégrées aux informations de combinatoire grammaticale. Dans ce dictionnaire, les collocations grammaticales et les collocations lexicales sont mêlées dans le même article. Elles sont regroupées implicitement par structure (et par analogie sémantique). Le BBI regroupe les collocatifs qui entretiennent une affinité sémantique. Les informations de valence (prépositions régies par la collocation, par exemple) sont mentionnées dans le BBI mais de façon non

systématique, même si ce dictionnaire prête une attention à ces aspects grammaticaux. Ce dictionnaire est basé sur le BBI dictionary of English word combinations (1997).

Il n'y a pas d'indicateur de catégorie dans la BBI. Le nom (ici : decision) est écrit en gras, en conséquence, il est visible. Sa définition est également présente à côté. Nous pouvons étudier le pivot dans différents contextes. Le dictionnaire nous montre quels mots peuvent antéposer ou postposer le mot qui nous intéresse. Pour l'exemple donné, il s'agit des verbes, des adjectifs et des prépositions qui sont tous énumérés l'un après l'autre. Nous pouvons également voir différentes structures syntaxiques dans lesquelles apparaît le mot étudié. De plus, il convient de mentionner qu'il n'y a pas malheureusement d'exemples d'usages.

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ Des collocations lexicales et grammaticales</li> <li>✓ Le regroupement par la structure</li> </ul>	<ul style="list-style-type: none"> <li>➤ Pas d'exemples d'usage</li> <li>➤ Manque de clarté</li> <li>➤ L'usage n'est pas vraiment simple</li> <li>➤ La disposition n'est pas vraiment bonne</li> </ul>

Pour le français, nous pouvons mentionner plusieurs dictionnaires :

- Le *Dictionnaire collocationnel du français général*, de Grobelak, L. (1990) (désormais DCFG). Varsovie, Państwowe Wydawnictwo Naukowe. Ce corpus est constitué d'articles de presse variés, romans, etc. Il contient 2 millions de mots.

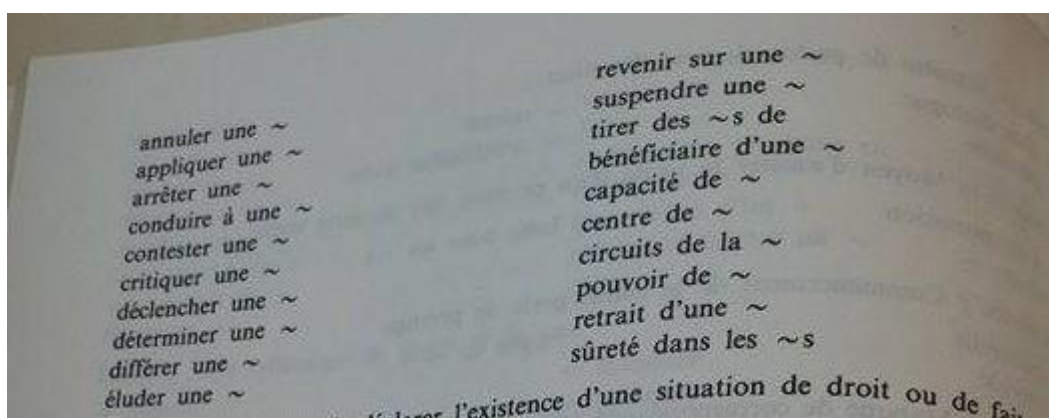
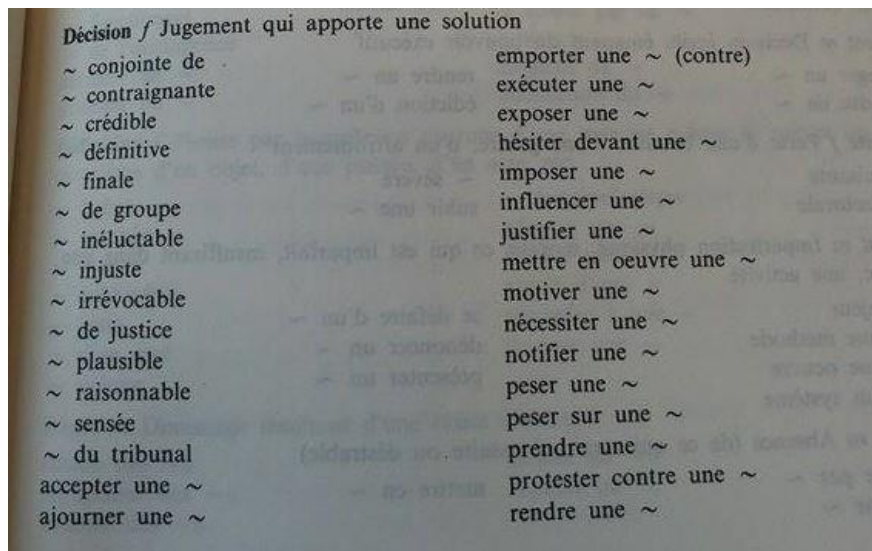


Figure 3 : Article de *décision* dans le DCFG

Il semble que ce soit le premier des dictionnaires de collocations du français. Le DCFG propose un codage redondant où la collocation est accessible parallèlement sous le collocatif et sous la base. Ainsi, la collocation *soulever l'enthousiasme* est doublement codée dans le dictionnaire : elle apparaît sous l'entrée de *soulever* dans la partie dédiée aux bases nominales, et dans la partie verbale sous *enthousiasme*. Ce mode de codage peut être pratique et intéressant, notamment de point de vue de l'utilisateur, mais il faut bien dire qu'il est peu économique pour un dictionnaire papier. Le nombre de bases n'est pas mentionné. Quant au nombre de collocations, il s'élève à 30 000 dont beaucoup apparaissent à la fois sous la base et sous le collocatif. Dans ce dictionnaire, le regroupement syntaxique est implicite et le classement sémantique est inexistant. Les mots sont regroupés par l'ordre alphabétique selon 3 parties :

I. substantif – substantif

II. verbe – substantif

III. adjectif – substantif

Nous trouvons dans ce dictionnaire la collocation avec l'emploi correct de la préposition. Grobelak est professeur du français. Il a trouvé qu' « une des plus grandes difficultés de l'apprenant [...] est de trouver le terme qui entre en collocation avec un autre terme, les associations collocationnelles n'étant pas – le plus souvent – identiques de langue à langue. Cela équivaut à trouver, selon la terminologie stylistique, le mot juste » (Grobelak, 1990 : 8). Ce dictionnaire est destiné aux apprenants du français ainsi qu'aux traducteurs. Le dictionnaire relève les collocations du français général (le mot *général* est utilisé pour le bien distinguer du mot *spécialisé*).

Le mot cherché (ici : *décision*) est écrit en gras. Nous trouvons sa définition à côté. Le tilde (~) avant ou après l'adjectif indique où le nom se trouve (cf. ~crédible → décision crédible). Le dictionnaire montre clairement si le pivot devrait être placé dans la structure donnée (devant, après ou au sein de la structure). De plus, il est indiqué si le mot pivot doit être employé au singulier ou au pluriel, ce qui est important du point de vue tant syntaxique que sémantique. Par ailleurs, les verbes et les adjectifs sont mélangés, il n'y a pas d'indicateur de catégorie.

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"><li>✓ L'usage facile</li><li>✓ La clarté</li><li>✓ L'emploi de la préposition qui va avec la collocation</li><li>✓ Le codage redondant</li></ul>	<ul style="list-style-type: none"><li>➤ Pas d'exemples d'usage</li><li>➤ Peu de collocations (juste 30 000)</li><li>➤ Adjectifs et verbes mélangés</li><li>➤ Déjà périmé</li></ul>

- Le *Dictionnaire des cooccurrences* (désormais DC), Beaugesne, J. (2001). Montréal, Guérin. Ce corpus a été établi à partir de sources très variés.

**FEUILLAGE** abondant, aéré, bruissant, caduc, clair, clairsemé, décoratif, délicat, dense, éclairci, épais, épineux, fin, foncé, frémissant, glauque, gracieux, gracie, houleux, impénétrable, informe, jaune, jauni, jaunissant, large, léger, lourd, luisant, lustré, luxuriant, maigre, mince, mouvant, naissant, obscur, persistant, poussiéreux, résistant, riche, robuste, sombre, soyeux, superbe, touffu, tremblant, vernis, vermissé, vert.  
*Arborer, offrir, posséder; porter, présenter; produire un (+ adj.).*  
 Le ~ frémit, tremble.

**FEUILLE** (*arbre plante*) allongée, ample, anguleuse, argentée, arrondie, brune, caduque, ciselée, composée, cotonneuse, découpée, dentée, dentelée, dorée, douce, dure, échancrée, étroite, fanée, ferme, filiforme, flétrie, fraîche, grasse, large, légère, longue, minuscule, molle, morte, naissante, nouvelle, odorante, ombrageuse, persistante, piquante, plissée, plate, pointue, raide, ridée, ronde, rouge, rougissante, rouillée, rousse, sèche, séchée, simple, sombre, souple, tachetée, tendre, tourbillonnante, tremblante, veloutée, velue, verdoyante, vernie, vermissée, verte.  
 Les ~s bougent, bruissent, frémissent, frissonnent, jaunissent, jonchent le sol, oscillent, susurrent, tombent, tourbillonnent, tournoient, tremblent, volent.  
 •(papier) blanche, courte, détachée, froissée, impeccable, lignée, longue, (im)maculée, mobile, propre, quadrillée, unie, vierge, volante.  
*Déchirer, froisser, (dé)plier une -; écrire sur une -.*

**COULEUR** affreuse, agressive, (im)altérable, ardente, assortie, belle, blafarde, blême, bonne, brillante, brutale, cadavérique, changeante, chargée, charmante, chatoyante, chaude, choquante, claire, contrastante, crayeuse, criarde, crue, cuivrée, défraîchie, dégradée, délavée, délicate, diaphane, discordante, discrète, disparate, (im)distincte, dorée, douce, douçonneuse, dure, éblouissante, éclatante, effacée, estompée, éteinte, étincelante, fade, fanée, fatiguée, féérique, flamboyante, floue, foncée, fondue, forte, fraîche, franche, frappante, froide, gaie, horrible, incertaine, incomparable, indéfinissable, indéterminée, inqualifiable, intense, intermédiaire, jeune, jolie, laide, laiteuse, lavée, légère, livide, louche, luisante, lumineuse, magnifique, mate, mélangée, merveilleuse, métallique, morte, neutre, nuancée, obscure, opaque, opposée, pâle, pâissante, passée, pauvre, (im)précise, prédominante, primaire, profonde, problématique, pure, ravissante, reposante, résistante, riche, secondaire, séduisante, sobre, sombre, soutenue, subtile, superbe, tendre, terne, tranchée, translucide, transparente, triste, unie, uniforme, vibrante, vieille, vilaine, violente, vive, voyante.  
*Avoir, posséder une - (+ adj.); être d'une - (+ adj.); adoucir, altérer, atténuer, (r)aviver, dégrader, éclaircir, faire ressortir, réchauffer, rehausser, relever, renforcer une -; amalgamer, assortir, combiner, discerner, harmoniser, mélanger, marier, varier des/les -s; changer de -.*  
 Des/les ~s contrastent, se heurtent, se marient, s'harmonisent, s'opposent.

**ARBRE** âgé, antique, bas, cavemieux, centenaire, chétif, colossal, court, creux, crevasse, cultivé, dégarni, dénudé, dépérissant, dépouillé, déraciné, desséché, dévasté, difforme, droit, dru, élané, élégant, élevé, émondé, énorme, épais, épineux, étendu, exotique, faible, (im)fertile, feuillu, fleuri, florissant, fort, foudroyé, fourchu, géant, gigantesque, grêle, gros, haut, immense, imposant, impressionnant, isolé, jeune, large, long, luxuriant, magnifique, maigre, majestueux, malade, massif, merveilleux, millénaire, mince, moribond, mort, moussu, nain, noueux, nu, ombreux, parfait, pointu, pourri, poussiéreux, (im)productif, puissant, rabougri, rachitique, ramifié, ratatiné, ravissant, robuste, rugueux, sauvage, sec, séculaire, solide, solitaire, sombre, somptueux, spacieux, stérile, superbe, svelte, tordu, tortueux, touffu, transplanté, trapu, vénérable, verdoyant, vert, vieil, vieillissant, vigoureux, vivant.  
*Abatte, arracher, coupe, cultiver, débiter, décapiter, décortiquer, dégarnir, dépouiller, déraciner, ébrancher, écimer; écorcer, effeuiller; élaguer, émonder, enlever, entailler, escalader; étêter, faire pousser, greffer, inciser, mutiler, (dé)planter, rabatre, redresser; taillader, tailler, transplanter, tronçonner, tuteurer un -; grimper, monter à un -; nicher dans un -; descendre d'un -; séjourner sous un -; grimper sur un -.*  
 Un ~ bourgeonne, croît, fleurit, frissonne, gémit, grandit, grossit, meurt, naît, penche, porte des fruits, pousse, prend racine, s'abat, se décompose, se défeuille, se dégarnit, se déploie, se développe, se dresse, s'effeuille, s'élève, se meurt, s'épanouit, se reproduit, s'évase, tremble, végète, (re)verdit, verdoie, vieillit, vit.

*Tiré du Dictionnaire des cooccurrences de Jacques Beauchesne*

Figure 4 : Article de *feuille* dans le DC

Le nombre de bases de ce dictionnaire s'élève à 4 200 bases, néanmoins le nombre de collocations n'est pas mentionné. Le regroupement syntaxique est implicite. Dans le DC, c'est le passage du romain à l'italique qui signale le changement de structure. Le DC adopte un classement mixte, c'est-à-dire, l'ordre alphabétique pour les collocatifs adjectivaux et synonymiques pour les constructions verbales. En cas de polysémie des bases, les entrées sont organisées en sous-entrées. Quant au corpus utilisé, ce sont des romans du 19e et 20e siècle, de la presse et des magazines.

Le mot recherché est écrit en gras et en majuscule (par contre nous n'y trouvons pas le genre du mot). Différentes acceptions d'un mot peuvent avoir différents collocatifs. Ceux-ci sont regroupés selon la signification du pivot. Les adjectifs qui vont avec le mot sont énumérés l'un après l'autre, il n'y a pas de lemmes, le mot est déjà accordé en nombre et en genre. Il y a également les verbes qui vont avec ce mot et des prépositions adaptés.

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ L'usage assez facile</li> <li>✓ La clarté</li> <li>✓ Un classement mixte</li> </ul>	<ul style="list-style-type: none"> <li>➤ Pas d'exemples d'usage</li> <li>➤ Le nombre de collocation pas mentionné</li> <li>➤ Le corpus utilisé (périmé déjà)</li> </ul>

- Le *Dictionnaire des combinaisons de mots : les synonymes en contexte* (désormais DCM) (sous la direction de Dominique Le Fur, 2007, Paris : Dictionnaire le Robert). Ce dictionnaire se compose de 160 000 combinaisons de mots classées par sens.

∞ **déception** + ADJECTIF

- amoureuse · sentimentale : *il s'est suicidé par déception sentimentale* · personnelle : *en dépit d'une immense déception personnelle*
- légère <sup>+nom</sup> · mineure · petite <sup>+nom</sup> · relative · passagère · ponctuelle
- évidente · manifeste · patente · perceptible · sensible · visible · complète · considérable · de taille · énorme · extrême · grande <sup>+nom</sup> · grosse <sup>+nom</sup> · immense · infinie · intense · lourde : *les résultats de cette année auront été une lourde déception* · partagée · profonde · vive <sup>+nom</sup> · véritable <sup>+nom</sup> · vraie <sup>+nom</sup>
- affreuse · amère · atroce · cruelle · douloureuse · horrible · terrible

∞ **déception** + VERBE

- être due à · naître de · résulter de · venir de · s'ensuivre
- percer : *une voix où perce la déception* · se manifester : *une certaine déception se manifeste parmi les employés* · s'exprimer
- être mêlée de · être teintée de : *leur déception est teintée d'amertume*
- être à la hauteur de l'attente / de l'espoir · être à la mesure de l'attente / de l'espoir : *la déception des perdants est à la mesure de leur espoir*
- attendre : *de grosses déceptions nous attendent*
- (plur.) s'accumuler · se multiplier

- adoucir · adoucir · calmer · contrebalancer · tempérer · relativiser
- épargner : *pour lui épargner une cruelle déception* · éviter · limiter
- cacher · dissimuler · masquer
- digérer · évacuer · oublier · ravalé · se remettre de · surmonter · effacer · noyer : *il noie sa déception dans l'alcool*

**décision** *nom fém.* (choix, parti)

∞ **décision** + ADJECTIF

- administrative · gouvernementale · judiciaire · préfectorale · présidentielle · budgétaire · économique · financière · politique · stratégique · etc.
- formelle · officielle · publique · souveraine · collective · collégiale · commune · unanime · individuelle · personnelle · unilatérale · prévisible
- nécessaire · capitale · cruciale · essentielle · importante · lourde de conséquences · majeure · définitive · ferme · finale · irréversible · irrévocable · sans appel
- prompte · rapide · immédiate · subite
- audacieuse · courageuse · hardie · héroïque · radicale · habile · judicieuse · bonne <sup>+nom</sup> · brillante · cohérente · éclairée · équilibrée · mesurée · motivée · pesée · pondérée · raisonnable · raisonnée · rationnelle · (mûrement /

Figure 5 : Article de *décision* dans le DCM

longuement) réfléchi - responsable - sage  
 • heureuse - salutaire - historique - sans précédent - claire - consensuelle

- favorable (à) : une décision favorable aux victimes - positive
- défavorable (à) : une décision défavorable au plaignant - négative : en cas de décision négative du ministre
- étonnante - étrange - inattendue - inhabituelle - surprenante
- tardive
- de principe - de pure forme
- compliquée - délicate - difficile (à prendre)
  - hasardeuse - périlleuse - risquée - douloureuse - grave - pénible
- irraisonnée - irréfléchie - irresponsable
  - contestée - controversée - critiquée - impopulaire - aberrante - absurde - ahurissante - choquante - consternante - contestable - inacceptable - incompréhensible - scandaleuse - inique - injuste - arbitraire - illégale - désastreuse - dramatique - fâcheuse - funeste - inadaptée - malheureuse - contre-productive - maladroite - mauvaise +<sup>nom</sup> - inopportune - intempestive - hâtive - précipitée - prématurée - ambiguë - contradictoire

∞ **décision** + VERBE

- découler de : cette décision découle en bonne partie des pressions internationales - faire suite à
- intervenir - prendre effet : cette décision prendra effet au mois de janvier
- appartenir à - être du ressort de : cette décision est du ressort de la médecine - incombent à - relever de : la décision relève de la seule compétence des autorités - revenir à
- conduire à - entraîner - provoquer : sa décision a provoqué une vive polémique - susciter
  - engager - mettre en jeu : une décision qui met en jeu un groupement d'intérêts économiques
- concerner - porter sur - affecter
- se fonder sur - se justifier - s'expliquer par
- faire grand bruit - soulever un tollé
- mûrir : le refus du chef de la diplomatie de lui parler au téléphone a précipité une décision mûrie de longue date
- passer inaperçue

∞ **VERBE** + **décision**

- attendre
- aboutir à - arriver à - déboucher sur - parvenir à
  - arracher - emporter - brusquer - forcer - imposer : le gouvernement tente d'imposer sa décision
- arrêter - prendre : prendre une décision sans hésiter ; la décision a été prise à l'unanimité - rendre : le recteur doit rendre sa décision dans la journée - maintenir - être conforté dans - appliquer - concrétiser : nouer les moyens pour concrétiser une décision - donner effet à - mettre à exécution - mettre en œuvre
- annoncer - donner - exprimer - formuler - informer de - notifier - rendre - rendre public - prononcer (Admin., Droit) : pour autant que la décision ait été prononcée en dernier ressort
- accepter - adopter - approuver - avaliser - entériner - voter : voter une décision à l'unanimité
  - adhérer à - appuyer - se rallier à - se ranger à - se soumettre à - s'incliner devant - applaudir à - saluer - confirmer - justifier - honorer - obéir à - respecter - se conformer à - se plier à
- expliquer - fonder : les arguments qui fondent ma décision - motiver
- guider - infléchir - influencer - influer sur - peser sur - avoir un droit de regard sur
- reconsidérer - regretter - revenir sur
- réagir à : l'accusé n'a pas réagi à la décision du tribunal
- ajourner<sup>nom</sup> - différer : la décision fut longuement différée - remettre (à plus tard) - reporter - repousser - retarder - geler : ils gèlent toute décision importante lors des campagnes électorales - suspendre
- (re)mettre en cause - remettre en question
  - contester - protester contre - s'inscrire en faux contre - aller à l'encontre de - entraver : rien ne peut entraver la décision du juge
- annuler - invalider (Admin., Droit)

<sup>1</sup> **déclaration** nom fém. (propos)

∞ **déclaration** + ADJECTIF

- gouvernementale - ministérielle - présidentielle - politique - sociale
- de principe - fondatrice - solennelle - sous serment - formelle - officielle : il s'est abstenu de toute déclaration officielle sur sa position dans ce conflit - publique - télévisée
- unilatérale - collective - commune
- de pure forme - spontanée : un témoin a fait une déclaration spontanée à la police - brève +<sup>nom</sup> - courte +<sup>nom</sup> - lapidaire - petite +<sup>nom</sup> - mesurée - modérée - prudente
- grande +<sup>nom</sup> - interminable - longue +<sup>nom</sup>
- (sur) nombreuses - répétées
- historique - importante

Figure 6 : Article de *décision* dans le DCM : la suite

Le DCM est un ouvrage volumineux. Il comporte un nombre impressionnant de collocations qui s'élève à 160 000 et qui a une nomenclature assez restreinte de 2600 noms. Dans ce dictionnaire, les différents types de structures sont introduits par un indicateur de catégorie, par exemple « VERBE + décision ». Il regroupe les collocatifs qui entretiennent une affinité sémantique. Par exemple, les collocatifs adjectivaux de *décision* qui mettent en jeu une dimension sociale sont regroupés sous la même puce dans le DCM *administrative, gouvernementale, judiciaire, etc.* Ce dictionnaire propose un ordre constant dans la présentation des groupements sémantiques, ce qui permet assez facilement d'en comprendre le sens. Il se concentre sur la position de l'adjectif épithète qui un point assez important dans l'apprentissage du français. Il y a aussi des exemples de l'utilisation des collocations qui sont importants pour la compréhension et pour le réemploi. Le dictionnaire est basé sur un corpus de textes journalistiques et littéraires de 500 millions de mots. Les traitements sémantiques dans ce dictionnaire sont assez fins et s'appuient essentiellement sur des regroupements synonymiques.



Le mot cherché est écrit en gras, à côté dans les parenthèses, nous trouverons ces synonymes et le genre du mot cherché. Ensuite, nous voyons les indicateurs de catégories qui facilitent vraiment l'usage (cf. Décision + ADJ).

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ L'usage assez facile</li> <li>✓ La clarté : bien organisé</li> <li>✓ Un grand corpus</li> <li>✓ Un indicateur de catégorie qui facilite l'usage</li> <li>✓ Les exemples</li> <li>✓ Les prépositions qui vont avec la collocation</li> </ul>	<ul style="list-style-type: none"> <li>➤ Le traitement sémantique – point assez faible</li> </ul>

- *Le Dictionnaire explicatif et combinatoire*, publié en 4 volumes entre 1984 et 1999 (désormais DEC)

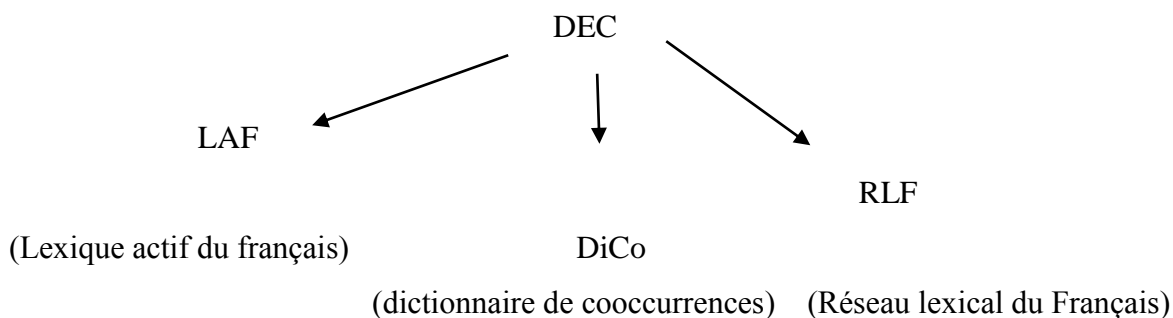
La lexicologie explicative et combinatoire décrit une méthode de construction d'articles de dictionnaire basée sur cette théorie. Cette méthode est indépendante des langues et elle permet ainsi de représenter n'importe quelle langue.

Le DEC est un dictionnaire linguistique, formel, avec une méthodologie de construction précise qui utilise un métalangage complexe. Déchiffrons maintenant l'abréviation du DEC = Dictionnaire explicatif et combinatoire : explicatif parce que le sens y est décomposé selon des principes rigoureux. Combinatoire, car la combinatoire syntaxique et lexicale y est décrite et c'est exactement la combinatoire qui est l'aspect le plus original dans ce dictionnaire. Les auteurs disent que le DEC est « un schéma idéalisé de dictionnaire de langue ». (Polguère, Mel'čuk, 2006). Chaque article suit un schéma préalablement défini, ce qui fait qu'ils ont une structuration rigide. Les auteurs soulignent que l'originalité du DEC réside dans le fait que c'est la première tentative d'élaboration d'un dictionnaire de langue formel qui soit ancré dans une théorie linguistique. Il faut ajouter que dans ce dictionnaire, toute la description lexicographique est explicite et que c'est le champ sémantique qui y domine. Ce dictionnaire ne décrit pas tous les phrasèmes de la même façon. Nous nous référons maintenant à la typologie mel'čukienne parlant des

*phrasèmes complets, quasi-phrasèmes et semi-phrasèmes*. C'est dans les semi-phrasèmes où nous trouvons les *collocations* du type ACCEPTER/DÉCLINER UNE INVITATION ou VIF INTÉRÊT) et qui sont décrits sous l'entrée d'un de leurs constituants - par ce qu'on appelle les fonctions lexicales. Le DEC fournit également des exemples négatifs qui sont souvent très pratiques pour soutenir la solution proposée. Par exemple, le verbe CHANGER a deux entrées dans ce dictionnaire. Les auteurs expliquent que c'est justifiée par rapport aux exemples du type *\*La mère a changé Colette de coiffure*<sup>TM</sup>. Cela va de même pour *la plus grande surprise de ce séjour et sa plus grande surprise*. En français, on ne dit pas - même si quelqu'un a eu une grande surprise pendant son séjour quelque part - *\*sa plus grande surprise de ce séjour* [il faut dire *sa plus grande surprise pendant ce séjour ou la plus grande surprise de son séjour ...* ] ; cet exemple négatif justifie la restriction qui accompagne le régime du lexème SURPRISEI.2

Le DEC n'est pas destiné au grand public, plutôt à un public éclairé. Il a été la source de l'inspiration pour d'autres dictionnaires, comme *le Lexique Actif du Français*. Le DEC connaît des adaptations pour le TAL, principalement le DiCo (développé par Alain Polguère)

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ L'envie de créer un dictionnaire idéalisé</li> <li>✓ Adapté à n'importe quelle langue</li> <li>✓ Très précis</li> <li>✓ Une nouveauté dans le domaine de la linguistique</li> <li>✓ La combinatoire comme le point très original dans ce dictionnaire</li> <li>✓ Une structuration rigide</li> <li>✓ Le corpus vaste</li> </ul>	<ul style="list-style-type: none"> <li>➤ Le dictionnaire est très, voire <b>trop</b> complexe</li> <li>➤ Il n'est pas adapté au grand public</li> <li>➤ L'usage <b>très</b> difficile</li> <li>➤ Quand même c'est un dictionnaire idéalisé</li> </ul>



Le DEC a engendré plusieurs dictionnaires « dérivés » plus grand public : LAF : Lexique Actif du Français : dictionnaire grand public de dérivations sémantiques et de collocations du français et DiCo (dictionnaire de cooccurrences) : version simplifiée et informatisée du DEC. Il y a encore le RLF (réseau lexical du Français – un dictionnaire virtuel).

### Les fonctions lexicales d'Igor Mel'čuk

Focalisons-nous maintenant sur les fonctions lexicales. D'après Igor Mel'čuk, chaque locuteur du langage, fait face au choix lexical. Comme exemple, nous pouvons dire que si le locuteur veut se référer au « véhicule dont le propriétaire est un particulier et qui l'utilise pour son usage personnel », il va dire tout simplement la « voiture ». Ce sont des choix lexicaux qui expliquent ce raisonnement. La théorie des fonctions lexicales est un ensemble d'outils fait pour décrire de manière systématique les types de relations lexicales que les lexèmes entretiennent entre eux. Il y a deux fonctions intéressantes pour nous dans cette théorie, sur lesquelles nous allons nous appuyer. Les voici : la fonction *Oper* et la fonction *Magn*.

La fonction *Oper* lie le mot-clé au verbe sémantiquement vide et le prend comme objet direct.

*Oper* – verbe support (*faire*). Verbe qui va avec *promenade (faire)* → *faire une promenade*

Quant à la fonction *Magn*, elle lie le mot-clé aux éléments qui expriment le plus haut degré du concept désigné par le mot-clé.

Magn – la collocation qui va avec : *remercier (infiniment)*

Igor Mel'čuk distingue plus de 50 fonctions lexicales.

Voici quelques exemples en français et en polonais :

➤ En Français :

Magn (pluie) = forte, violente < torrentielle, diluvienne

Magn (argument) = fort, convaincant

Oper<sub>1</sub> (voyage) = effectuer, faire, entreprendre

Oper<sub>1</sub>(excuses) = présenter (~à N)

➤ En polonais:

Magn (deszcz 'pluie') = silny 'forte', ulewny 'torrentiel'

Magn (argument 'argument') = silny 'fort', mocny 'puissant'

IncepOper<sub>1</sub> (podróż 'voyage') = wybrać się [w~] 'aller', wyruszyć [w~] 'partir'

Oper<sub>1</sub> (przeprosiny 'excuses') = -

Pour le polonais, Oper<sub>1</sub> valable pour « podróż » 'voyage' n'existe pas. La valeur de IncepOper<sub>1</sub> a été présentée comme un exemple alternatif. Il n'y a pas non plus Oper<sub>1</sub> valable pour « przeprosiny ». Au lieu de cela, le verbe *przepraszać* est utilisé (Wanner, 1996)

La modélisation des fonctions lexicales [=FL] permet de voir deux types de relations lexicales distinctes : les dérivations sémantiques et les relations base-collocatif, autrement appelées, les *collocations*. La base de la collocation désigne à la fois un vrai mot et à une sorte de métalangage. (Fontenelle, 1992) Parmi les FL existantes, seules les FL syntagmatiques (par opposition aux FL paradigmatiques comme la synonymie, l'hyponymie, l'antonymie, etc.) permettent de modéliser les relations des collocations (Grossmann & Tutin, 2003 :12).

## Les collocations dans les logiciels d'aide à la rédaction

*Antidote* (société *Druide*) est un programme qui aide à écrire sans fautes. Il dispose de dictionnaires avec recherche puissante et des guides linguistiques. Par ailleurs, les guides et les dictionnaires interagissent et aident à explorer les aspects linguistiques. Le programme dispose encore des correcteurs d'orthographe, de syntaxe et du style qui „analysent” le langage et aident à corriger un texte. Il présente également les collocations en contexte et en corpus et exploite des informations de fréquence qui sont importants de point de vue didactique (Tutin, 2007). *Antidote* contient plus de 125 000 mots, dont 13 000 noms propres, incluant les noms des États et des personnages célèbres. Il répertorie 46 000 locutions courantes et spécialisées, dont chacune est accompagnée de sa définition et d'exemples. Le dictionnaire des cooccurrences dresse un inventaire complet des associations de mots. Au Total, 900 000 combinaisons de mots (*Antidote*, *posologie*). Focalisons-nous maintenant sur le dictionnaire des cooccurrences.

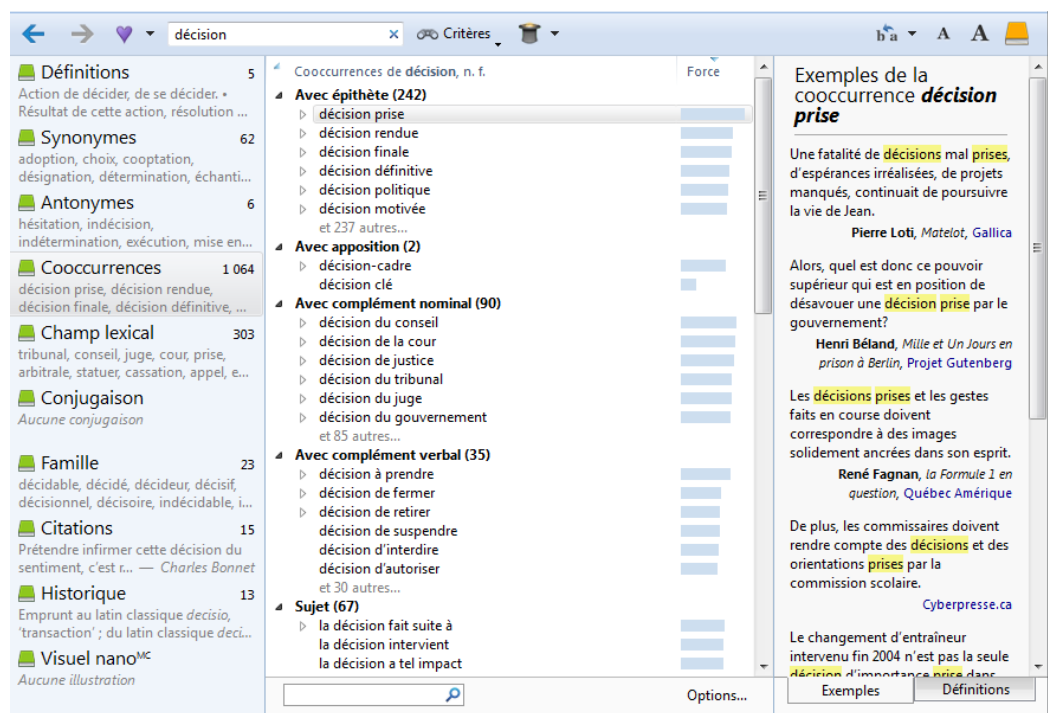


Figure 7 : Dictionnaire des cooccurrences d'Antidote : *décision*

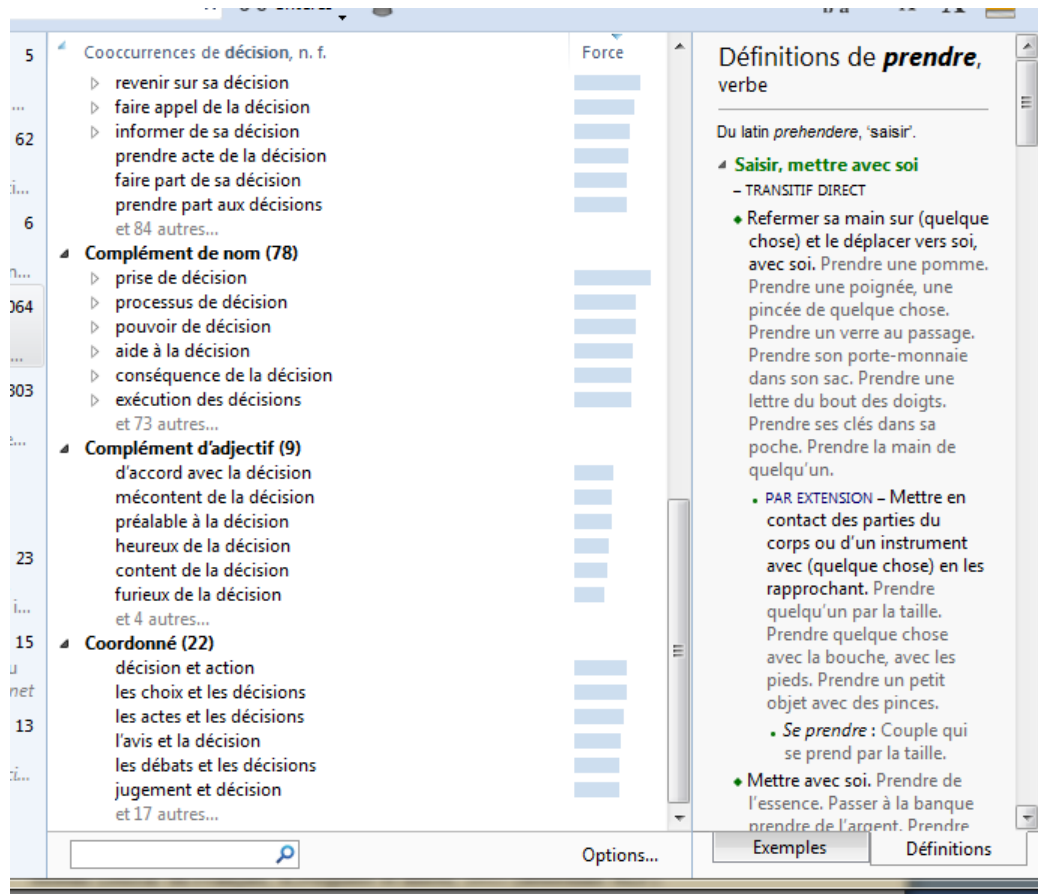


Figure 8 : Dictionnaire des cooccurrences d'Antidote, la suite : *décision*

Ce dictionnaire présente l'ensemble *des combinaisons lexicales les plus marquantes* du mot en vedette qui sont classées selon :

- le contexte syntaxique
- la force de cooccurrence
- si c'est possible, selon les divers sens du mot en vedette.

Par *combinaison lexicale marquante*, nous comprenons deux ou plusieurs mots (la *base* et les *cooccurents*) dont l'association dans la langue est forte, souvent spontanée et fréquente.

Pour expliquer le concept, nous pouvons nous demander, comment, par exemple, qualifier *une décision*. Ce peut être : (une) *décision prise, rendue, finale, définitive, politique*, etc. Le dictionnaire des cooccurrences fournit une liste des associations statistiquement les plus fortes, dans tous les contextes syntaxiques intéressants.

Les cooccurrences du mot en vedette apparaissent en liste verticale. Ce qui est intéressant à noter, c'est le fait qu'une cooccurrence peut être fléchie au féminin ou au pluriel si elle se présente plus souvent ainsi dans la langue. Ensuite, un trait bleu après chaque cooccurrence indique sa force statistique relative : si le trait est plus long, cela signifie que la cooccurrence est plus forte par rapport aux autres cooccurrences dans le même contexte. La force dépend avant tout de la fréquence, mais aussi de la singularité de l'association des termes.

Les cooccurrences sont classées par contexte syntaxique : *épithète, sujet, complément*, etc. ; le nombre et la nature des contextes dépendent de la catégorie du mot en vedette. Le nombre de cooccurrences pour chaque contexte apparaît entre parenthèses après le nom du contexte. Dans le cas où un mot possède plusieurs sens, les cooccurrences sont regroupées par sens et ils sont affichés en vert.

La sélection d'une cooccurrence donne accès, dans le panneau de droite, à des exemples qui illustrent la cooccurrence en contexte réel. Dans chaque exemple, les mots de la cooccurrence sont surlignés. Chaque exemple est présenté avec un nom de l'auteur et le titre de l'ouvrage. En bleu, il y a également le nom du corpus source utilisé. C'est un lien hypertexte qui permet d'ouvrir le site Web du corpus dans un moteur de recherche.

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ l'usage numérique, facile</li> <li>✓ beaucoup d'exemples</li> <li>✓ bonne organisation</li> <li>✓ destiné au grand public</li> <li>✓ efficace</li> <li>✓ un logiciel complexe et riche de possibilités</li> <li>✓ une cooccurrence peut être fléchie au féminin ou au pluriel si elle se présente plus souvent ainsi dans la langue</li> <li>✓ la force de la collocation</li> </ul>	<ul style="list-style-type: none"> <li>➤ les propriétés syntaxiques ne sont pas bien expliquées</li> </ul>

- le Robert Correcteur

Le *Robert* est l'un des meilleurs dictionnaires en langue française. Maintenant, nous trouvons un projet du « Robert correcteur ». C'est un correcteur orthographique de nouvelle génération. Il analyse le texte, en d'autres termes, il analyse le sens des phrases et souligne chaque type d'erreur d'une couleur différente et en donne l'explication. Il vérifie également la ponctuation. Ce dictionnaire s'appuie sur la technologie du correcteur ProLexis de Diagonal qui est utilisé par les plus grandes maisons d'éditions et institutions. Il intègre les contenus de références des éditions Le Robert.

Le Robert Correcteur se compose entre autres de 8 dictionnaires de référence qui contiennent 250 000 définitions, 35 000 noms propres, 3 millions de synonymes et de contraires, 18 000 locutions, 1,4 million de combinaison de mots, 17 000 citations françaises et étrangères, 8 000 proverbes, 10 000 verbes conjugués.

Focalisons-nous avant tout sur le dictionnaire des combinaisons. Si nous comparons Antidote et le Robert Correcteur, c'est l'homologue du dictionnaire des cooccurrences d'Antidote. Regardons le mot "décision" dans ce dictionnaire. Il y a une représentation intéressante. Il y a d'abord une vue graphique du mot employé avec les adjectifs. Quand l'adjectif est plus gros, cela signifie que la fréquence de cette combinaison est importante. À droite, il y a la liste des combinaisons. Quand les combinaisons sont trop nombreuses pour y être présentées graphiquement, il y a un petit ascenseur qui permet de passer de série en série. Quand nous voudrions voir des combinaisons, nous cliquons sur l'un des adjectifs. Après, nous pouvons voir d'autres emplois, par exemple, afficher les verbes ayant « décision » pour complément direct ou les combinaisons où le mot cherché (ici : décision) est complément d'un autre nom ou dans lesquelles „décision” ou un autre mot cherché est le sujet. Les possibilités sont vastes.





Figure 9 : Dictionnaire des combinaisons de Robert Correcteur : *décision*



Figure 10 : Dictionnaire des combinaisons de Robert Correcteur : la suite : *décision*

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ Interface efficace</li> <li>✓ Agréable à utiliser</li> <li>✓ Des explications claires</li> <li>✓ Représentation innovante</li> </ul>	<ul style="list-style-type: none"> <li>➤ les propriétés syntaxiques ne sont pas bien expliquées</li> </ul>

Quant aux logiciels, nous nous sommes permis de résumer des observations dans une grille :

<b>ANTIDOTE</b>	<b>ROBERT CORRECTEUR</b>
<ul style="list-style-type: none"> <li>- Les deux programmes sont assez complémentaires</li> <li>- La fenêtre de correction, autrement dit, les explications de fautes: il faut les trouver par un clic</li> <li>- Il permet de corriger les fautes dans le texte directement dans la fenêtre de correction</li> <li>- Il possède une sorte d'anticorps – nous pouvons l'attacher à l'email, il lance la correction automatiquement avant envoi, corrige un texte dans un programme de mail.</li> </ul> <p><b>Dictionnaire :</b></p> <ul style="list-style-type: none"> <li>- Antidote a développé ses cooccurrences sur des citations.</li> <li>- Les collocations classées selon le contexte syntaxique (épithète, sujet, complément, etc), la force de cooccurrence et les divers sens du mot en vedette.</li> <li>- Une cooccurrence peut être fléchie</li> </ul>	<ul style="list-style-type: none"> <li>- Les deux programmes sont assez complémentaires</li> <li>- La fenêtre de correction : les explications de fautes sont affichées clairement en permanence</li> <li>- Il ne permet pas de corriger les fautes directement dans la fenêtre de correction</li> <li>- Il n'y a pas de cette fonction ici</li> <li>- <b>Dictionnaire des combinaisons :</b> a d'abord été construit à partir de travaux de linguistes qui ont développé les exemples</li> <li>- Le dictionnaire des combinaisons du Robert est très original dans sa présentation (graphes, etc)</li> <li>- Les collocations classées selon le contexte syntaxique</li> <li>- Il y a la liste des combinaisons (le mot cherché en tant que le sujet ou complément d'un nom, etc.)</li> </ul>

au féminin ou au pluriel si elle se présente plus souvent ainsi dans la langue	
--	--

### Le dictionnaire des collocations en polonais

En polonais, il y a un projet qui aborde de façon plus ou moins largement le problème de collocations. C'est qui est intéressant, c'est le fait qu'en anglais ou en français, on trouve des dictionnaires des collocations. Cependant, en polonais, ce sujet n'est pas vraiment bien abordé. Et pourtant le phénomène de la collocation existe dans toutes les langues et leur réalisation est spécifique à l'intérieur de chaque langue. En polonais, depuis quelques années, le développement des corpus slaves est assez rapide. À partir de 1994 quand les travaux sur le Grand Corpus Tchèque ont débuté, des corpus référentiels des autres langues slaves sont apparus. Des corpus croates, polonais, russes, slovaques et slovènes sont disponibles sur Internet. Il faut quand même souligner que la morphologie slave est un grand défi pour les nouvelles technologies. Nous pouvons même dire que c'est un défi plus complexe que la morphologie des langues romanes (Lewandowska-Tomaszczyk, 2012). Ainsi, en polonais, il y a des corpus qui permettent de rechercher des collocations. Quant aux dictionnaires, nous ne trouvons qu'un seul qui s'intéresse aux associations typiques entre des mots, ainsi, à la notion de *collocation* : le Dictionnaire SŁOWNIK DOBREGO STYLU, CZYLI WYRAZY, KTÓRE SIĘ LUBIĄ (Mirosław Bańko, 2006). Ce dictionnaire contient environ 5000 bases, dans la plupart de cas, ce sont des substantifs. Il énumère les associations typiques entre elles. Au total, il propose 70 000 associations. Le dictionnaire fournit des réponses telles que : *co robi parlament ? (qu'est-ce que le parlement fait-il ?) (powołuje kogoś (désigne qqn), przyjmuje uchwałę (adopte une résolution), uchwala coś (adopte qqch), zatwierdza (approuve projet de loi)), co można zrobić z paszportem ? (qu'est-ce que peut-on faire avec un passeport ?) (dostać (recevoir), uzyskać (obtenir), wyrobić sobie (faire une demande))*. Ce dictionnaire n'explique pas la signification des mots. Nous n'y trouvons pas de définitions ni informations grammaticales. Son but est de rassembler les associations de mots les plus naturelles et les plus fréquentes en polonais. Les données linguistiques présentées par le dictionnaire sont issues du Corpus de la langue polonaise PWN.

L'auteur postule que le dictionnaire est compréhensible pour tout le monde, quels que soient l'âge ou la connaissance (linguistique). Le dictionnaire est divisé en 3 parties principales – selon les parties du discours : *noms*, *verbes* et *adjectifs*. Dans chaque partie, il y a des collocations indiquées dans la langue polonaise. Pour des mots spécifiques – qui sont classés selon l'ordre alphabétique – il y a d'autres mots adaptés qui ensemble, forment des collocations.

À titre d'exemple, prenons le mot « chêne » (*dąb* en polonais). Pour le mot *dąb*, il y a des mots qui y sont adaptés : *kilkusetletni* (séculaire), *okazały* (majestueux) , *pomnikowy* (monumental) , *potężny* (puissant), *rozłożysty* (branchu) , *stary* (vieux) , *stuletni* (centenaire), ...

Ainsi, il y a un certain nombre de termes et c'est au lecteur de choisir quel terme lui convient le mieux dans le contexte donné.

The screenshot shows a dictionary entry for the Polish word 'aktywność'. The entry is highlighted with a red box. Above the main entry, there are several lines of text listing related terms and their meanings, such as 'popularny', 'dramatyczny', 'scen polskich', 'gra', 'kreuje jakąś postać', 'odtwarza jakąś postać', 'wciela się w jakąś rolę', 'występuje'. The main entry for 'aktywność' lists various adjectives and nouns associated with it, such as 'duża', 'gorączkowa', 'nadmierna', 'niezwykła', 'podwyższona', 'wielka', 'wysoka', 'wzmoczona', 'znacząca', 'zwiększona', 'mała', 'niska', 'słaba', 'ludzka', 'obywatelska', 'społeczna', 'artystyczna', 'intelektualna', 'twórcza', 'umysłowa', 'zawodowa', 'życiowa', 'dyplomatyczna', 'ekonomiczna', 'gospodarcza', 'kulturalna', 'polityczna', 'biologiczna', 'chemiczna', 'fizyczna', 'metaboliczna', 'wulkaniczna', 'ruchowa', 'seksualna', 'człowieka', 'państwa', 'rządu', 'społeczeństwa', 'komórki', 'mózgu', 'słońca', 'brak', 'obniżenie', 'ograniczenie', 'osłabienie', 'spadek', 'utrata', 'wahanie', 'zahamowanie', 'zanik', 'zmniejszenie', 'pobudzanie', 'wzrost', 'zwiększenie', 'wplywać na', 'czyjaś lub czegoś', 'pobudzać', 'stymulować', 'wyzwalać', 'zwiększać', 'hamować', 'zmniejszać', 'przejawiać', 'wykazywać', 'wyzwała się', 'wzrasta', 'zwiększa się', 'zmniejsza się', 'przejawia się w czymś'. Below the main entry, there are three more entries: 'alarm', 'album', and 'alejja', each with its own list of related terms and meanings.

Figure 11 : Article de *aktywność* (activité) dans le *Słownik Dobrego Stylu*

Points FORTS (+)	Points FAIBLES (-)
<ul style="list-style-type: none"> <li>✓ Les associations de mots les plus naturelles</li> <li>✓ Les associations les plus fréquentes</li> <li>✓ Compréhensible pour tout le monde</li> <li>✓ Une aide pour ceux qui s'occupent</li> </ul>	<ul style="list-style-type: none"> <li>➤ Pas d'informations grammaticales</li> <li>➤ Pas vraiment bien lisible, les informations stockés l'une à côté de l'autre</li> </ul>

de la traduction ✓ Le premier dictionnaire des collocations en polonais ( !)	
--	--

Nous avons observé que dans les dictionnaires étudiés, la collocation est en général codée à partir de la base, qui est, autrement dit, l'élément stable de la collocation. Par exemple, la collocation *prendre la décision* est codée dans l'article de *décision* dans le *Dictionnaire des cooccurrences* de Beauchesne. Ce mode de codage n'est pas problématique quand la distinction entre base et collocatif est claire, mais comme c'est souvent le cas de certaines expressions de structures N prép N, l'analyse peut être souvent plus complexe. Pour les objectifs de cette étude, nous avons choisi un sous ensemble de type N prep N. Pour clarifier le concept, prenons comme l'exemple l'expression *meute de chiens*. Tout dépend si nous voulons considérer *meute* comme une sorte de quantifieur, où *chien* peut être considéré comme la base, ou bien nous disons que *meute* qui est précisée sémantiquement par *chien* devient la base. Dans des cas qui sont ambigus, nous soutenons le point de vue que les dictionnaires de collocations devraient proposer une double entrée pour ces éléments. Rappelons que le *Dictionnaire Collocationnel du Français Général* de Grobelak propose un codage redondant où la collocation est accessible parallèlement sous le collocatif et sous la base, ce qui est également le cas dans le logiciel Antidote. Ainsi, la collocation *prendre la décision* est doublement codée : elle apparaît sous l'entrée de *prendre* dans la partie dédiée aux bases nominales, et dans la partie verbale sous *décision*. Cette façon de faire semble être peu économique pour le format papier, mais peut-être il s'avère que c'est une bonne idée pour un dictionnaire en ligne.

Le traitement syntaxique présent dans des dictionnaires de collocations est très succinct. Il porte avant tout sur les types de constituants (néanmoins, cette information n'est pas présente dans tous les dictionnaires, à savoir, elle est absente dans le DC). Dans le OCDSE, il y a le constituant occupé par le collocatif qui précise l'ordre des mots. Dans d'autres dictionnaires de collocations (le DC ou le DCM), les structures des collocations ne sont pas trop diversifiées. Finalement, nous pouvons apercevoir que les informations syntaxiques ne sont pas traitées suffisamment et ne permettent pas toujours facilement d'employer le dictionnaire de façon adaptée en encodage. Nous pouvons ajouter que dans le DC, l'ordre des adjectifs (anté- ou postposé) n'est pas indiqué.

Quant au traitement sémantique, il est absent dans certains dictionnaires comme le DC. Un traitement par rapprochement synonymique est proposé dans le DCM ou l'OCDSE où les collocatifs sont regroupés par analogie. Le DEC indique la structure syntaxique dans laquelle le mot s'insère. Il propose un ensemble de notions de base et la description fine et rigoureuse des relations lexicales fournies qui aident à mieux comprendre le fonctionnement des structures lexicales en français. Il propose une présentation formelle. Quand nous parlons des *collocations*, l'analyse peut s'aborder selon deux axes : d'abord, nous pouvons nous concentrer sur l'étude des collocatifs pour une lexie donnée. Par exemple, étudier les collocatifs de la lexie *décision*. Après nous pouvons mettre en relief les expressions typiques d'une fonction lexicale. Nous trouvons que le DEC est un outil très riche, mais sa présentation est complexe et sa consultation nécessite une formation de haut niveau.

### ***1.5. Synthèse***

Depuis quelques années, l'intérêt sur des études sur la phraséologie a permis de déboucher sur des produits lexicographiques vraiment innovants. On a ainsi vu dans les pays anglo-saxons, à côté des classiques dictionnaires d'idiomes, des dictionnaires de collocations dès les années 80. La plupart de ces dictionnaires sont destinés avant tout aux apprenants des langues étrangères. Dans le premier chapitre, nous avons fait un petit inventaire des dictionnaires des collocations dans 3 langues étrangères, à savoir l'anglais, le français et le polonais.

Focalisons – nous sur les propriétés syntaxiques dans les dictionnaires présentés ci-dessus. Dans l'OCDSE, nous pouvons observer des mots qui apparaissent avec le collocatif. Ils sont regroupés selon leur catégorie grammaticale à laquelle ils appartiennent et sont introduits par un indicateur de catégorie (cf. Verbe+decision). Dans le BBI, nous pouvons regarder le mot étudié dans différents contextes. Ce dictionnaire montre quels mots peuvent être anté- ou postposés au mot cherché. Dans le DCFG, nous voyons si le mot doit être placé devant, après ou au sein de la structure. Il marque également si le mot cherché doit être employé au singulier ou au pluriel, ce qui est utile de point de vue syntaxique et sémantique. Dans le DC, nous trouvons différentes acceptions d'un mot qui peut avoir différents collocatifs (cf. Feuille (arbre, plante))

ou (papier)). Les adjectifs qui vont avec le mot sont énumérés l'un après l'autre, il n'y a pas de lemmes, le mot est déjà accordé en nombre et en genre. Il y a également les verbes qui vont avec ce mot et des prépositions adaptés. Dans le DCM, il y a un indicateur de catégorie qui introduit différents types de structures (cf. verbe + décision). Ce dictionnaire propose un ordre constant dans la présentation des groupements sémantiques, ce qui permet assez facilement d'en comprendre le sens. Il se concentre sur la position de l'adjectif épithète. Quant au DEC, c'est un dico idéalisé avec un métalangage spécifique. Il est basé sur les fonctions lexicales d'Igor Mel'čuk. La théorie de FL<sup>4</sup> est un ensemble d'outils fait pour décrire les types de relations lexicales que les lexèmes entretiennent entre eux. Quant au dictionnaire polonais, il est peu spécialisé, il montre les associations qui sonnent les plus naturelles pour un locuteur natif. Il y a des adjectifs qui sont anté- ou postposés au mot cherché, déjà fléchi à la forme cherchée. Dans le Robert Correcteur, il y a une représentation graphique et une liste des combinaisons où nous pouvons voir des emplois et des combinaisons du mot cherché.

---

<sup>4</sup> FL – fonctions lexicales

## **Partie 2**

-

### **Méthodologie et problématique**



## Chapitre 2. Problématique : propriétés syntaxiques de collocations

Notre travail s'est inscrit dans la continuité du cadre du projet ANR Termith (Terminologie et Indexation de Textes en sciences Humaines) avec les travaux du laboratoire LIDILEM (Laboratoire de Linguistique et Didactique des Langues Etrangères et Maternelles) de l'Université Grenoble Alpes (Stendhal 3 de Grenoble) sur un lexique de base transdisciplinaire des écrits scientifiques. Nous sommes partis d'une liste de collocations déjà extraite dans le cadre de ce projet. Notre étude se limite aux collocations de type N prép N.

La méthode consiste à extraire les propriétés syntaxiques des corpus, puis d'en repérer leur fonctionnement. Dans ce chapitre, nous présenterons les corpus utilisés dans le cadre du projet Termith, et l'outil d'extraction utilisé, le Lexicoscope. Nous finirons par un inventaire des propriétés syntaxiques examinées à partir d'un état de l'art sur les structures N1 de N2.

La problématique du présent mémoire concerne les propriétés syntaxiques de collocations qui est un sujet peu traité en linguistique. La majorité des travaux se concentrent sur l'aspect lexical, la syntaxe y est négligée. C'est pourquoi nous souhaitons nous intéresser à ce problème qui jusqu'alors a été peu exploré. Dans ce qui suit, nous voudrions présenter une modélisation de ces associations préférables. Il faudrait présenter une description sémantique fine de la collocation (définition) et de ses propriétés syntaxiques. Prenons comme l'exemple la collocation *enquête de terrain*. Dans ce cas-là, la définition doit être accompagnée de données syntaxiques mentionnant le caractère inséparable du substantif et de substantif (N de N). De plus, il faudrait ajouter que *terrain* dans ce cas-là ne possède pas d'article. Notre motivation pour les propriétés syntaxiques de collocations est liée à la fois aux exigences dans les écrits scientifiques, et aux attentes des apprenants. L'écrit scientifique exige l'usage d'un langage spécifique, autrement dit, du langage de spécialité. Quant aux étudiants, les apprenants ont beaucoup de souci pour reconstituer une collocation. Néanmoins, les collocations sont présentes dans les langues et il faut savoir les utiliser correctement. Certaines collocations acceptent une permutation lexicale à l'aide de quasi-synonymes, comme *avancer/émettre une hypothèse*. Par contre, un polonophone pourrait faire des fautes en disant \* « poser une hypothèse » car il s'appuie naturellement sur les modèles de sa langue maternelle. Ainsi, nous voyons que les propriétés de la combinatoire syntaxique suggèrent de ne pas séparer leurs éléments

constitutifs afin de conserver le sens de ces associations lexicales. Les relations entre les éléments constitutifs des collocations sont à prendre en considération. Nous allons démontrer dans la partie suivante qu'il existe une différence entre *champ d'étude* et *champ de l'étude* qui ne doit pas être évidente pour les apprenants.

## 2. *Corpus et outil d'extraction*

Dans le cadre de notre projet, nous avons essayé de repérer des constructions syntaxiques dans l'écrit scientifique. Nous avons recouru au corpus Scientext qui est un corpus d'écrits scientifiques.

Notre approche linguistique s'inscrit dans le courant de la *linguistique de corpus*. Les études sur les corpus permettent d'observer les phénomènes lexicaux dans leur contexte naturel. À partir de corpus, nous pouvons analyser l'usage de certaines expressions (ici : *collocation*), pour déterminer si cela présente des difficultés particulières et donc, ce qui nous permettra à contribuer à l'élaboration d'un dictionnaire en ligne. Les corpus peuvent servir à formuler des hypothèses (corpus-driven) ou à les vérifier (corpus-based) (Tongini-Bonelli,2001).

Commençons par définir la notion de corpus. John Sinclair, un des fondateurs de la « Corpus Linguistics », l'a définie comme ceci :

« *A corpus is a collection of pieces of language that are selected according to explicit linguistic criteria in order to be used as a sample of the language* » (Sinclair 1996 : 4).

Nous pouvons alors dire que le corpus est un ensemble de textes (oraux ou écrits) recueillis d'après quelques critères. Les corpus sont utilisés dans plusieurs domaines : études littéraires, linguistiques, etc. La branche de la linguistique s'occupant des corpus est liée aux développements des systèmes informatiques, en particulier à la constitution de bases de données textuelles. Ainsi, il s'agit des technologies informatiques qui ont pour but d'aider les traducteurs dans leurs travaux ou les apprenants d'une langue étrangère en vue d'apprendre certaines constructions qui ne sont pas faciles à prévoir (cf. *les collocations*). Nous différencions plusieurs types de corpus (cf. corpus spécialisé qui contient une langue de spécialiste, caractéristique pour un domaine ou le corpus de référence qui fournit des informations à propos d'une langue).

John Sinclair (2004) évoque plusieurs critères concernant les corpus. Avant tout, un corpus est censé avoir la quantité, c'est-à-dire, il doit être grand, fiable et doit contenir un nombre important de données. Ensuite, il doit être authentique et contenir des informations naturelles – c'est la qualité d'un corpus. Le corpus devrait être codé en norme ASCII (c'est la norme de base, la plus utilisée), c'est la simplicité qui en exige. Finalement, le corpus doit être documenté, c'est-à-dire, il doit avoir des informations comprenant les sources, les métadonnées, etc. L'utilité des corpus est incontestable. Ils donnent la possibilité d'observer des mots dans le contexte, permettent de voir si les expressions fonctionnent toujours dans une langue, montrent des mots spécialisés (d'une langue scientifique par exemple) et le vocabulaire d'un domaine spécifique.

En français, il existe plusieurs corpus de langue française, mais distinguons avant tout les bases d'écrits scientifiques puisque c'est le sujet qui nous intéresse le plus.

### **3. *Le corpus Scientext des sciences humaines***

*Scientext* s'inscrit dans le cadre de l'ANR, intitulé « Corpus et outils pour étudier le positionnement et le raisonnement de l'auteur dans les écrits scientifiques ». Il comprend plusieurs corpus : des textes scientifiques en Français (5 millions de mots, 205 textes), les textes scientifiques en anglais (34 millions de mots, 3381 articles), des écrits universitaires d'étudiants d'anglais langue étrangère (1, 1 million de mots, 300 textes d'apprenants de l'anglais) et des évaluations de Communications (502 évaluations). Le corpus *Scientext* a été basé sur le classement des revues, à savoir le classement de l'AERES (l'Agence de l'évaluation de la recherche et de l'enseignement supérieur). Il englobe plusieurs disciplines : linguistique, économie, psychologie, sociologie, biologiemedecine, chimie, informatique. Les sous-genres représentés sont les articles, les résumés d'articles, les communications, les thèses, les rapports, les HDR. Il a été annoté structurellement avec le standard XML TEI Lite, ce qui a permis de reconnaître les parties textuelles (introduction, conclusions, notes de bas de page). Par ailleurs, chaque forme comporte un lemme, une partie de discours, des traits morphosyntaxiques (nombre, genre, temps, mode) et est reliée à d'autres mots par des dépendances syntaxiques. Grâce à cela, nous pouvons rechercher des constructions plus générales, ce qui permet de se baser pas seulement sur l'aspect lexical, mais également sur l'aspect grammatical. C'est une ressource de textes scientifiques mis à dispositions pour des linguistes,

épistémologues, spécialistes de la recherche d'information, scientifiques. Il est disponible en ligne : <http://scientext.mshalpes.fr/scientext12/?lang=fr&corpus=sctexts>.

Ce corpus propose plusieurs types d'annotations : d'abord, les en-têtes, une annotation de la structure des articles, l'annotation morpho-syntaxique. Puis, l'annotation des en-têtes fournit des informations sur le texte + annotation syntaxique (l'origine de l'écrit, le traitement effectué, le type d'écrit, son type) et l'expertise de l'auteur. Ensuite, la structure du texte est annotée d'une manière semi-automatique, ce qui permet de déterminer les parties textuelles (résumé, introduction, notes, références etc.).

#### ***4. L'outil de traitement de corpus utilisé : Lexicoscope***

Par ailleurs, dans le cadre de ce mémoire, nous effectuerons une recherche en utilisant l'outil de traitement de corpus le *Lexicoscope* qui fut créé par Olivier Kraif et Sascha Diwersy, d'abord initialement développé dans le cadre du projet Emolex. Emolex est un projet franco-allemand sur le lexique des émotions dans 5 langues européennes. (Kraif, Diwersy, 2012)

Il est accessible en ligne : <http://dip01.u-grenoble3.fr/~kraifo/lexicoscope/index.php>

Cet outil a été employé dans plusieurs travaux : le projet franco-allemand Emolex (2009 – 2012), pour l'extraction de pivots complexes pour l'exploration de la combinatoire du lexique : une étude dans le champ des noms d'affect (Kraif, Tutin, Diwersy, 2014), projet Phraseotexte (2015-2016) : l'étude de la phraséologie stéréotypée du Roman français contemporain (2009 – 2012), Phraseorom (2016-2019) : la phraséologie du Roman.

*Le Lexicoscope* est un outil gratuit d'exploration de textes qui cherche l'occurrence d'un mot (ou d'une expression) et qui fournit la liste des occurrences avec leurs contextes. Il est centré sur l'étude des profils combinatoires des unités lexicales. En travaillant avec ce programme, nous pouvons choisir une des 5 langues européennes : anglais, français, espagnol, russe et allemand. Cet outil est aussi un concordancier pour les arbres syntaxiques (treebanks) qui permet d'analyser facilement les propriétés syntaxiques des collocations.

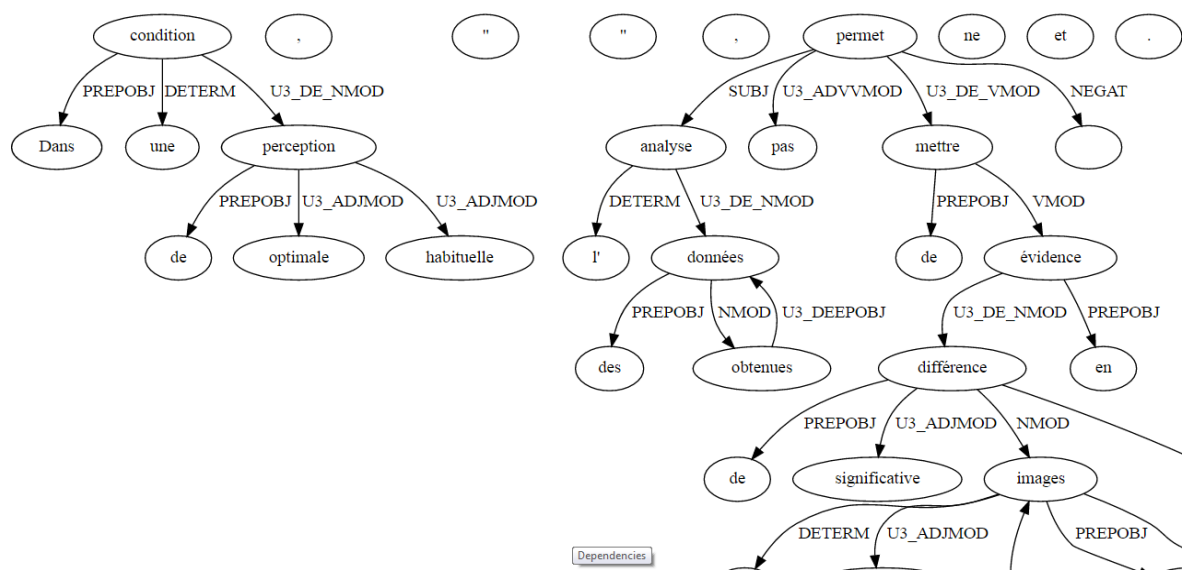


Figure 13 : l'arbre dans le programme Lexicoscope : analyse syntaxique de la phrase *Dans une condition de perception optimale, « habituelle », l'analyse des données obtenues ne permet pas de mettre en évidence de différence significative entre les images structurées et les images déstructurées (psychologie).*

Les collocations binaires simples peuvent être extraites à l'aide d'une relation syntaxique et des mesures d'association. Le Lexicoscope contient plusieurs corpus. Il possède un choix comprenant des corpus monolingues et parallèles.

Pour notre recherche qui consiste à repérer et à décrire les propriétés syntaxiques dans l'écrit scientifique, nous avons utilisé un corpus scientifique de plusieurs domaines développé dans le cadre du projet Termith : économie, sociologie, sciences politiques, psychologie, linguistiques, sciences de l'éducation, sciences de l'informatique, histoire, géographie, anthropologie constitué de 500 textes. La taille totale du corpus est de presque 5 millions de mots.

Brièvement décrit, le Lexicoscope permet d'analyser les concordances et les cooccurrences des mots. Il existe deux manières de soumettre une requête : soit dans le mode dit « simple », soit dans le mode « avancé ». Ce dernier nous permet de faire des analyses plus complexes. Par contre, son usage, n'est pas trop facile. Ainsi, la requête « simple » est une première recherche qui permet de voir, entre autres, le nombre d'occurrences d'un mot donné sans relation syntaxique identifiée. La requête « avancée »

est une deuxième possibilité qui permet d'effectuer une recherche beaucoup plus complexe. Nous pouvons y établir une relation syntaxique d'un mot avec un autre.

Afin de voir le contexte dans lequel le collocatif apparaît avec le mot pivot, il faut cliquer sur le collocatif. Ainsi, nous obtiendrons tous les contextes. Les informations telles que l'auteur, le titre de la revue dans laquelle se trouve une expression cherchée, l'édition et l'année de publication y sont incluses. Il est possible de vérifier le contexte plus large d'une collocation où nous pouvons observer le comportement d'un mot dans une phrase.

Nous pouvons aussi visualiser les résultats sous la forme d'un arbre syntaxique. De même, nous pouvons y observer une hiérarchie ainsi que les relations syntaxiques de chaque élément.

Hormis l'arbre syntaxique, il est possible d'observer l'annotation syntaxique de la phrase sous forme d'un tableau qui spécifie l'ordre des mots, la forme, le lemme, la catégorie, les traits et la relation syntaxique.

L'utilisation du Lexicoscope est assez bien développée. Elle permet de repérer les mots les plus fréquents et d'en analyser les concordances et les cooccurrences. Avant de se lancer dans la recherche, il est quand même nécessaire de bien connaître le programme et d'en apprendre les formules qui nous permettront d'effectuer les recherches les plus complexes.

## **5. 5 d'analyse proposée**

Dans le cadre de ce mémoire, nous allons nous focaliser sur les expressions de type *N préposition N*. C'est un groupe particulier, une forme de construction assez productive.

Certaines expressions peuvent être trouvées dans les écrits scientifiques. Il existe plusieurs types de structures de type N1 de N2:

- 1) Les structures où N1 est un nom quantitatif : nom + série (*une série de questions, une série d'enquêtes, etc.*) ou ensemble (*un ensemble de facteurs, un ensemble de propriétés, etc.*), avalanche (*une avalanche de questions, etc.*).

Si nous prenons l'exemple suivant : *une avalanche de questions*, ici le mot *avalanche*, est considéré comme quantificateur de type *beaucoup* et signale une quantité restreinte. Ces collocations diverses sont autorisées avec *avalanche* : *une avalanche de questions, problèmes, cadeaux, surprises, etc.* Néanmoins, il semble qu'il existe une restriction concernant les noms non comptables, par exemple : *\*une avalanche d'eau ?* *Avalanche de* semble permettre d'exprimer la quantité s'il s'agit d' « objets » (abstraites ou concrets). La motivation sémantique de telles expressions est toujours bien vivace (Tutin, 2004).

- 2) Les structures où N1 est un nom qualitatif : *une sorte de* ou *un type de*.
- 3) Les structures appelées N de qualité (Larrivée, 1997 : 134) : *nom de Dieu, ton idiot de frère, un drôle de type, etc.*

Il y a bien d'autres structures – cf. Grossman, Tutin (2016)

Dans la partie analyse, nous allons examiner certains exemples de type N prép N. Cette construction est une construction fréquente en français. C'est un groupe particulier, vu sa nature polyfonctionnelle, en même temps également très intéressant pour l'étude et la classification des collocations dans le domaine du discours général et spécialisé. Dans notre étude, nous avons essayé d'observer la caractérisation sémantique de N1 et N2 (si elle est prédicative ou non ainsi que leur relation mutuelle), le rôle joué par le PREP N2 (présence ou absence d'un argument syntaxique), le rôle joué par la préposition, les spécifications du nombre grammatical de N1 et N2, ), les relations entre N1 et N2 et N2 par rapport à N1 et la présence ou l'absence de déterminants derrière la N2. Comme le montreront les exemples suivants dans la partie analyse, parfois, il est difficile d'établir une règle générale, une tendance qui soit liée à l'absence de la détermination avant N2.

Il est important de savoir que les différentes caractéristiques impliquées interagissent et ne peuvent donc pas être considérées isolément.

Il y a des cas intéressants où le N1 joue un rôle de classifieur qui indique la perspective dans laquelle le sens de N2 doit être considéré. Prenons comme l'exemple *cadre d'analyse* qui est le cadre scientifique dans lequel une analyse est faite. Les noms, comme *champ, domaine* ont en général un espace (métaphorique).

En premier lieu, nous observons plusieurs paramètres, qui nous paraissent être essentiels pour notre recherche. Les voici :

1. Le N1 : est-ce qu'il se met au pluriel ou est-il toujours singulier?
2. Le N2 : est-ce qu'il peut être au pluriel ?
3. Quels sont les déterminants les plus fréquents devant le N1 ?
4. Est-ce qu'il y a des contraintes concernant le Prep avant le N2 ?
5. Le rôle joué par la préposition, les déterminants et le nombre grammatical: cela se rapporte en particulier à la présence ou l'absence d'un déterminant avant N2, et le type et l'ordre des déterminants (déterminant par exemple défini avec N1, déterminant indéfini avec N2, etc.): cela est une question très complexe, qui mérite d'être étudiée amplement, mais nous avons trouvé important de mentionner ce point ici.
6. Est-ce qu'il y a une possibilité d'insertion des adjectifs avant ou après le N1 aussi qu'avant ou qu'après le N2 ?

Le rôle de la préposition dans la structure N prép N, et surtout le rôle de la préposition *de* permet expliquer certaines relations entre le N1 et le N2. Nous nous limitons plutôt aux constructions qui sont utiles de point de vue des études transdisciplinaires.

Suivant le raisonnement de Benetti (1995) and Bartning (1996) cité par A.Tutin et F.Grossman (2015), nous pouvons observer que :

1. N1 est un nom prédicatif et N2 est son objet direct.

Dans cette catégorie, le N1 exprime le processus mental. Le N1 se réfère au processus mental pendant que le N2 représente l'objet et syntaxiquement, il est le second argument de N1. Le N2 se réfère à l'objet de l'analyse ou de l'évaluation; Il spécifie le champ d'application d'analyse ou d'évaluation, par exemple analyse de l'*activité* ou des *activités*, analyse de l'*effet* ou l'analyse des *effets*.

2. Il y a des cas où le N2 peut être au singulier ou au pluriel, mais le N1 est rarement utilisé au pluriel. C'est le cas par exemple de complexité du processus (des processus) (*\*Les complexités du processus*).



3. Le N2 dans les écrits scientifiques est souvent un nom d'idéalité, autrement dit, le N2 exprime un jugement, il est généralement un nom abstrait singulier, appartenant à des noms d'idéalité ou d'activité, comme, c'est par exemple le cas de *critère de sélection*.

En vue de l'analyse des propriétés syntaxiques des collocations, nous avons élaboré une grille qui nous permet de mieux apercevoir certaines affinités. Partants de la liste des collocations déjà effectuée dans le cadre du programme Termith, nous avons choisi quelques exemples des collocations de type N prép N. Nous avons choisi des collocations les plus fréquentes et les plus représentatives pour l'écrit scientifique. Pour effectuer cette tâche, nous avons fait la recherche dans le Lexicoscope. Pour chaque collocation, nous avons vérifié le nombre d'occurrences, c'est-à-dire, combien de fois cette collocation apparaît dans le corpus. Ensuite, nous avons vu les disciplines dans lesquelles telle ou telle collocation apparaît le plus souvent. Les disciplines sont classées selon la fréquence. Par exemple, pour la collocation *analyse des données*, nous voyons qu'elle apparaît le plus souvent en psychologie. Après, nous présentons un exemple d'usage de chaque collocation.

Voici cette liste :

1. **Critère de sélection** (15 cooccurrences)
2. **Analyse des données** (41 cooccurrences)
3. **Augmentation du nombre** (27 cooccurrences)
4. **Enquête de terrain** (38 cooccurrences)
5. **Champ d'étude** (20 cooccurrences)
6. **Objectif de recherche** (30 cooccurrences)
7. **Groupe de travail** (80 cooccurrences)
8. **Objet d'étude** (71 cooccurrences)
9. **Période d'observation** (34 cooccurrences)
10. **Point de départ** (71 cooccurrences)
11. **Processus de construction** (48 cooccurrences)
12. **Programme de recherche** (74 cooccurrences)
13. **Recherche en science** (67 cooccurrences)
14. **Réponse à une question** (15 cooccurrences)
15. **Résultat de l'étude** (43 cooccurrences)
16. **Recueil de données** (30 cooccurrences)
17. **Principe d'organisation** (18 cooccurrences)
18. **Point d'appui** (28 cooccurrences)
19. **Partie de l'article** (23 cooccurrences)
20. **Outil de communication** (30 cooccurrences)
21. **Objet d'analyse** (30 cooccurrences)
22. **Mode d'action** (38 cooccurrences)

- 23. Ensemble de pratique (34 cooccurrences)
- 24. Diffusion d'information (28 cooccurrences)
- 25. Valeur de référence ( 67 cooccurrences)

Dans le Lexicoscope, nous avons effectué nos recherches dans le mode « avancé », à chaque fois nous affinons nos requêtes pour que le programme nous sorte des informations demandées. Présentons deux requêtes :

**<l=analyse,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<l=donnée,c=NOUN,#3>::(PREP OBJ,3,2) (U3\_DE\_NMOD,1,3)**

L – lemme (une forme canonique). Dans ce cas-là, le programme peut nous sortir le nom soit au pluriel, soit au singulier <l=analyse> → *analyse ou analyses* ;

C - catégorie (dans ce cas-là : *analyse* est un nom, *donnée* – un nom, *de* – une préposition. À tour de rôle, nous avons un codage affecté à chaque catégorie. Ici pour nom, on code „*noun*”, pour la préposition „*prep*”, etc.) ;

Entre parenthèses, il y a des relations de dépendances. Ce sont les étiquettes des relations syntaxiques. À titre d'exemple **PREPOBJ,3,2** est une préposition qui est en relation avec l'objet.


**<w=analyses,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<l=donnée,c=NOUN,#3>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)**

Dans ce cas-là, nous vérifions si le N1 (ici : *analyse*) peut être au pluriel. Pour effectuer cette tâche, nous avons mis un « w » qui est une forme graphique. Alors, <w=analyses> → *analyse seulement au pluriel* ; Voir annexe 1 pour regarder les requêtes utilisées pendant notre recherche. Présentons maintenant la grille d'analyse et les éléments étudiés :

- a. *Sans contraintes* : D'abord, nous prenons une collocation telle qu'elle est. Nous ne mettons aucune restrictions dans la requête et nous voyons le nombre total

- d'occurrences dans notre corpus. Dans ce cas-là, le N1 peut être au singulier ou au pluriel, de même pour le N2 et la préposition *de* peut varier.
- N1 pluriel* : Nous vérifions si le N1 peut être mis au pluriel. Dans ce cas-là, le N1 est codé au pluriel et le reste des éléments en tant que lemmes.
  - N2 pluriel* : Nous codons N1 comme singulier et N2 comme pluriel. Nous vérifions si la collocation existe quand le N1 est singulier et le N2 est au pluriel.
  - DET+N1 singulier* : Nous vérifions quel déterminant va avec la collocation quand le N1 est au singulier.
  - DET+N1 pluriel* : De même, pour le déterminant avant le N1 au pluriel
  - N2 avec déterminant* : Nous vérifions si le N2 a un déterminant.
  - N1 pluriel N2 pluriel* : Nous vérifions si toute la collocation peut apparaître au pluriel (N1 et N2), cf. *Analyses de(s) données*
  - (ADJ1) N1 (ADJ1) de N2* : Nous vérifions si le N1 peut être suivi par l'adjectif (anté- ou postposé)
  - N1 de Adj2 N2* : Nous vérifions s'il y a des adjectifs qui vont avec le N2
  - N1 de N2 Adj2* : Nous observons s'il y a des adjectifs que l'on peut mettre à la fin de la collocation.

Voici la grille élaborée pour la collocation *analyse des données* :

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
<p>Sans contraintes</p> <p>Nous prenons une collocation et nous la vérifions dans le corpus sans aucune contrainte</p>	<p>Analyse(s) de (DET) donnée(s)</p> <p>&lt;1=analyse,c=NOUN,#1&gt;&amp;&amp;&lt;1=de,c=PRE P,#2&gt;&amp;&amp;&lt;1=donnée,c=NOUN,#3&gt;::(PREP OBJ,3,2) (U3_DE_NMOD,1,3)</p>  <p>Voici la requête que l'on a utilisée dans le Lexicoscope</p>	<p>41</p> <p>Le programme sort le nombre total d'occurrences pour cette collocation (sans contraintes)</p>	<p>6</p> <p>Psychologie Sciences éducation Sciences Info Sciences Politiques Sociologie Géographie</p> <p>Nous voyons les domaines dans lesquels la collocation apparaît (classés selon la fréquence)</p>	<p>L'approche quantitative s'appuie sur des données d'enquête et utilise l'<b>analyse de données</b> pour , d'une part , identifier les caractéristiques des internautes ayant des amis virtuels , notamment en termes de vie sociale réelle, et , d'autre part, caractériser ces amitiés virtuelles .</p> <p>Un exemple : la collocation dans l'usage</p>	<p>author =&gt; Virginia Lethiais Title =&gt; Réseaux 164 L'auteur et l'ouvrage d'où la collocation a été tirée</p>

<p>N1 pluriel</p> <p>Nous vérifions si la collocation existe quand le N1 est pluriel</p>	<p><b>Analyses</b> de(s) donnée(s)</p> <p>&lt;w=analyses,c=NOU N,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;::(PREPOBJ,3,2) (U3_DE_NMOD,1,3)</p>	<p>5</p>	<p>3</p> <p>Psychologie Sciences education Sciences politiques</p>	<p>Deux analyses des données sur les comportements pro-puis antisociaux ont validé deux échelles comprenant chacune une structure factorielle en 4 dimensions .</p>	<p>author =&gt; V. Leoni title =&gt; Le travail humain</p>
<p>N2 pluriel</p> <p>Nous vérifions si la collocation existe quand le N2 est pluriel</p>	<p>Analyse de(s) <b>données</b></p> <p>&lt;w=analyse,c=NOUN,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;w=données,c=NOUN,#3&gt;::(PREPOBJ,3,2) (U3_DE_NMOD,1,3)</p>	<p>33</p>	<p>6</p> <p>Psychologie Sciences education Sciences info Sciences politiques Sociologie Géographie</p>	<p>L'analyse des données s' est ensuite focalisée sur le lien entre le fait de percevoir un aspect bloquant / stimulant et la réalisation ou non d'un diagnostic de stress .</p>	<p>author =&gt; I. Hansez title =&gt; Le travail humain</p>
<p>DET + N1 singulier</p> <p>Nous vérifions quel déterminant va avec la collocation quand le N1 est au singulier</p>	<p><b>DET analyse</b> de(s) donnée(s)</p> <p>&lt;w=analyse,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;&amp;&amp;&lt;c=DET,#4&gt;::(PREPOBJ,3,2) (*,1,4)(*,1,3)</p>	<p>24</p> <p>23 : le 1: un</p> <p>Dans cet exemple, il y a 24 occurrences au total : 23 pour l'article défini (le) et 1 pour l'article indéfini un.</p>	<p>6</p> <p>Psychologie Sciences politiques Sciences info Géographie Sciences éducation Sociologie</p>	<p>Il faut donc reprendre <b>l'analyse de données</b> commerciales dans des créneaux technologiques de niveau plus élevé pour voir dans quelle mesure il convient de fortement nuancer les conclusions qui pouvaient être tirées de l' analyse de ces secteurs intensifs en travail .</p>	<p>author =&gt; Claude Grasland title =&gt; L'Espace géographique Vol. 39</p>
<p>DET + N1 pluriel</p> <p>De même, pour le déterminant avant le N1 au pluriel</p>	<p><b>DET analyses</b> de(s) donnée(s)</p> <p>&lt;w=analyses,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;&amp;&amp;&lt;c=DET,#4&gt;::(PREPOBJ,3,2) (*,1,4)(*,1,3)</p>	<p>1</p> <p>1: leurs</p>	<p>1</p> <p>Psychologie Sciences Education Sciences politiques</p>	<p>Par <b>leurs analyses</b> secondaires <b>des données</b> Eurobaromètre mais aussi par leur collaboration directe au programme, les chercheurs universitaires ont largement contribué à lui assurer ce quasi-monopole d' expertise sur l' analyse empirique des attitudes européennes .</p>	<p>author =&gt; Philippe Aldrin title =&gt; Politix 89</p>

<p>N2 avec déterminant</p> <p>Nous vérifions si le N2 a des déterminants ou pas</p>	<p>Analyse(s) de <b>DET</b> donnée(s)</p> <p>&lt;l=analyse,c=NOUN,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;&amp;&amp;&lt;c=DET,#4&gt;::(PREPOBJ,3,2) (U3_DE_NMOD,1,3) (DETERM,3,4)</p>	<p>37</p> <p>35 : des 1 : ces 1 : diverses</p>	<p>6</p> <p>Psychologie Sciences education Sciences politiques Géographie Sociologie Sciences info</p>	<p><b>L'analyse de ces données</b> indique une interaction significative entre la congruence frontière syllabique (...)</p>	<p>-</p>
<p>N1 pluriel N2 pluriel</p> <p>Nous vérifions si toute la collocation peut apparaître au pluriel</p>	<p><b>Analyses</b> de(s) <b>données</b></p> <p>&lt;w=analyses,c=NOUN,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;w=données,c=NOUN,#3&gt;::(PREPOBJ,3,2) (U3_DE_NMOD,1,3)</p>	<p>5</p>	<p>3</p> <p>Psychologie Sciences education Sciences politiques</p>	<p>Deux types d'analyses des données ont été réalisées : la première relève d' une analyse terminologique et thématique , afin de dégager les thématiques et termes récurrents pour caractériser les perceptions et vécus dans les espaces de travail.</p>	<p>author =&gt; V. Moulin title =&gt; Le travail humain</p>
<p>N1 ADJ1 de N2</p> <p>Nous vérifions si le N1 peut être suivi par l'adjectif (anté- ou postposé)</p>	<p>Analyse ADJ de(s) donnée(s)</p> <p>&lt;l=analyse,#1&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;&amp;&amp;&lt;c=ADJ,#4&gt;::(PREPOBJ,3,2) (.*,1,4)(.*,1,3)</p>	<p>9</p> <p>Analyse <b>qualitative</b> Analyse <b>secondaires</b> Analyse <b>statistique</b> Analyse <b>inductive</b> Analyse <b>structurale</b></p>	<p>4</p> <p>Sciences éducation Psychologie Sciences politiques Sciences info</p>	<p>L' étude tente d'une part de caractériser ces images identitaires en utilisant une méthodologie <b>d'analyse inductive des données brutes.</b></p>	<p>-</p>
<p>N1 de Adj2 N2</p> <p>Nous vérifions s'il y a des adjectifs qui vont avec le N2</p>	<p>Analyse de(s) ADJ donnée(s)</p> <p>&lt;l=analyse,c=NOUN,#1&gt;&amp;&amp;&lt;c=ADJ,#4&gt;&amp;&amp;&lt;l=de,c=PREP,#2&gt;&amp;&amp;&lt;l=donnée,c=NOUN,#3&gt;::(PREPOBJ,3,2) (U3_DE_NMOD,1,3) (.*,2,4)</p>	<p>0</p>	<p>0</p>	<p>-</p>	<p>-</p>
<p>N1 de N2 Adj2</p> <p>Nous vérifions s'il y a des</p>	<p>Analyse de(s) donnée(s) ADJ</p>	<p>11</p> <p>Analyse des données</p>	<p>5</p> <p>Sciences education Géographie</p>	<p>Il faut donc reprendre <b>l'analyse des données commerciales</b> dans</p>	<p>author =&gt; Claude Grasland title =&gt;</p>

adjectifs qui vont avec toute la collocation.	<l=analyse,c=NOUN,#1>&&<c=ADJ,#4>&&<l=de,c=PREP,#2>&&<l=donnée,c=NOUN,#3>:(PREPOBJ,3,2)(U3_DE_NMOD,1,3)(. *,3,4)	<b>brutes</b> Analyse des données <b>agrégées</b> Analyse des données <b>temporelles</b> Analyse des données <b>disponibles</b> Analyse des données <b>contextuelles</b> Analyse des données <b>commerciales</b> Analyse des données <b>bibliographiques</b>	Psychologie Sciences info Sciences politiques	des créneaux technologiques de niveau plus élevé pour voir dans quelle mesure il convient de fortement nuancer les conclusions qui pouvaient être tirées de l'analyse de ces secteurs intensifs en travail .	L'Espace géographique Vol. 39
---	--	--	---	--	----------------------------------

Après avoir analysé cet exemple donné, nous pouvons apercevoir que la forme typique de cette collocation est **l'analyse des données**. C'est une collocation fréquente dans l'écrit scientifique (41 cooccurrences).

Ainsi, la grille nous a permis de voir directement les résultats. Le déterminant qui apparaît avec le N1 au singulier (ici : *analyse*) est un article défini (*l'* – la forme élidée de *le*). Quant au pluriel, en prenant en considération notre corpus, nous n'avons pas trouvé d'exemples pertinents (seulement 1 occurrence) concernant l'article *leurs*. Selon les résultats, nous pouvons insérer plusieurs adjectifs dans cette collocation. Dans notre corpus, nous avons observé l'insertion de ces adjectifs avant ou après le N1 : *qualitative, secondaire, statistique, inductive, structurale*. Quant aux adjectifs se référant au N2, nous en avons observé 7 dans notre corpus : *analyse des données brutes/agrégées/temporelles/disponibles/contextuelles/commerciales/bibliographique*.

Comme on la voit, la préposition *de* dans la collocation *analyse de données* peut subir quelques changements : la préposition *de* se lie à l'article défini de N2 (ici : *les données*) ce qui nous donne l'article défini contracté et la forme *des* : (*de + les = des*). Il y a également l'article démonstratif *ce* qui apparaît au pluriel dans notre corpus : *l'analyse de ces données* et l'adjectif *divers* : *l'analyse de diverses données*. On peut également observer dans quelle mesure cette collocation est partagée par toutes les disciplines des sciences humaines. Il y a 6 domaines dans lesquelles elle apparaît, à

savoir : Psychologie, Sciences éducation, Sciences Info, Sciences Politiques, Sociologie, Géographie.

Finalement, nous pouvons remarquer quelques contraintes :

1. La forme typique se montre au singulier: *l'analyse des données*. La forme au pluriel n'est pas vraiment présente dans notre corpus (1 cooccurrence pour *leurs analyses...*).
2. Le déterminant qui apparaît le plus souvent devant le N1 est un article défini (*le*), plus précisément, sa forme élidée *l'*.
3. Le déterminant avant le N2 est un article défini contracté : il se fait par la liaison de la préposition *de* et de l'article défini du N2 (ici : *les données*). Ainsi, *de+les = des*. D'où vient la forme typique *analyse des données*.
4. Il y a des possibilités d'insérer des adjectifs après le N1 et après le N2.
5. Le N2 est toujours au pluriel (ici : *données*).

Exemple	Remarque	N1	déterminant	N2
<b>Analyse dess données</b>	-	- peu de contraintes	- article défini contracté	- est toujours au pluriel  -un nom abstrait

L'outil de traitement de corpus – le *Lexicoscope* nous a permis d'avoir les recherches rapides et efficaces. Nous avons choisi 25 collocations les plus représentatives et les plus fréquents pour l'écrit scientifique (le nombre de cooccurrences oscillait entre 15 (cf. *critère de sélection*) et 74 (cf. *programme de recherche*). Il faut dire que le nombre de collocations choisis était peu important mais suffisant pour effectuer notre tâche dans le cadre de ce mémoire. La grille élaborée pendant cette étude a permis de voir directement les résultats.

## **Partie 3**

-

### **Analyse et proposition de la modélisation**



### Chapitre 3. Analyse des propriétés syntaxiques des collocations et pistes de codage

La méthode et la grille d'analyse ont été appliquées à 25 collocations. Nous avons fait les mêmes grilles pour chaque collocation choisie (voir annexe 2). L'analyse de ces collocations avait pour objectif d'observer quelles modifications elles peuvent subir. Nous avons aussi essayé de trouver des ressemblances entre elles au niveau syntaxique. Cela nous a permis de voir si le N1 ou le N2 peuvent prendre la forme du pluriel, de quels déterminants elles peuvent être accompagnées, ou bien de vérifier si la préposition « *de* » (dans la structure N1 prép N2) peut être en lien avec un déterminant, et, si oui, avec lequel. Une telle étude nous a donné la possibilité de saisir des nuances au niveau sémantique entre différentes structures syntaxiques de la même collocation. L'ensemble des données analysées nous facilitera, en effet, d'effectuer notre tâche finale consistant à proposer un modèle d'un dictionnaire.

Nous examinerons ici quelques exemples pour présenter la méthodologie employée. Nous présenterons ensuite les résultats de l'étude sur le corpus des 25 collocations examinées.

La méthodologie présentée sur un exemple choisi :

#### **Forme typique: le champ d'étude**

#### **Lemme : champ d'étude**

Nous avons observé plusieurs résultats notables. La forme typique de cette collocation est **le champ d'étude**. La forme *le champ de l'étude* est également possible. Par contre, elle n'est pas si fréquente dans notre corpus (juste 3 occurrences contre 20). Il y a une petite différence entre *le champ d'étude* et *le champ de l'étude*. Nous pouvons dire que *champ d'étude* et *champ de l'étude* sont très proches, même si le premier a une plus large extension (se référant à un champ disciplinaire ou sous-disciplinaire), et l'autre (le deuxième) se réfère au domaine d'une étude spécifique.

La différence sémantique porte sur l'absence du déterminant de N2. Dans la construction N prép N, *champ* est souvent associé au nom *étude*. Le sens de *champ d'étude* et *le champ de l'étude* sont naturellement différents en raison de déterminants qui accompagnent chacun d'eux. La collocation *champ d'étude* apparaît dans les phrases

comme *notre champ d'étude* („Seul le cas de chômeurs indemnisés par le régime d'assurance chômage est considéré : les chômeurs non indemnisés ou qui perçoivent des allocations du type ASS (allocation spécifique de solidarité) ou RMI (revenu minimum d'insertion) ne font pas partie de **notre champ d'étude**”) ou quand l'auteur veut clarifier son domaine de recherche: *le champ des études culturelles* („L'école de Los Angeles se situe sur un autre versant de la géographie, soit du côté des approches qualitatives, littéraires pour une bonne part, et dans **le champ des études culturelles** d'orientation postmoderne”, G.Sénechal, *Annales de géographie*)

Il y a également des changements concernant le deuxième nom de la collocation. Ainsi, nous pouvons dire *le champ des études* (2 occurrences) ou *champ de cette étude* (1 occurrence). La collocation n'existe pas au pluriel (les deux noms ne peuvent pas se mettre au pluriel \* *Les champs des études*). Par contre, soit le N1 soit le N2 peut être au pluriel à condition que le 2<sup>e</sup> nom soit au singulier (*Le champ des études, les champs d'étude*).

Quant aux adjectifs, il y a des possibilités d'insertion avant et après le premier nom (ici : *champ*). Dans notre corpus, nous avons observé l'apparition de ces adjectifs (*Champ académique d'étude, Premier champ d'étude, Deuxième champ d'étude, Ce nouveau champ d'étude*). Nous n'avons rien observé quant à l'insertion de l'adjectif devant le N2.

Dans notre corpus, il y a 3 adjectifs qui vont avec cette collocation, à savoir :

- *Champ des études culturelles* ou l'adjectif au pluriel „culturelles” est lié au N2;
- *Champ des études urbaines* ou l'adjectif „urbaine” se réfère également au N2;
- *Champ d'étude suffisamment vaste* ou l'adjectif „vaste” qualifie le N1, et pas le N2.

Il y a 6 disciplines dans lesquelles cette collocation apparaît, à savoir : Sciences éducation, Géographie, Économie, Sociologie, Sciences info, Histoire (classés selon la fréquence).

Pour résumer, nous pouvons remarquer quelques contraintes :

1. La forme typique est au singulier: *le champ d'étude*.
2. Le déterminant qui apparaît le plus souvent devant le N1 est un article défini (*le*).

3. Il y a des possibilités d'insérer les déterminants avant et après le N1

Exemple	Remarque	N1	déterminant	N2
<b>Champ d'étude</b>	Le N1 classe le N2	- le N1 appartient à la classe sémantique de noms d'espace	- possibilité d'insérer un DET devant le N2	- le N2 est en général au singulier

Le N1 (ici : *champ*) est un nom d'espace qui localise et spécifie le N2. Autrement dit, nous voyons que le champ décrit de quelle étude il est question.

**Forme typique: les critères de sélection**

**Lemme : critère de sélection**

Les résultats de la présente étude démontrent que la forme typique de collocation **critère de sélection** est la forme au pluriel *les critères de sélection*. Le déterminant avant le nom *critère* est un article défini *le*. Ainsi, quant au pluriel, le déterminant qui apparaît le plus souvent devant le nom *critères* est un article défini *les* (*les critères de sélection*). Si nous voulons vérifier dans notre corpus s'il est possible de modéliser cette collocation, nous nous apercevons que les variations possibles portent sur le singulier et le pluriel de premier nom (ici : *critère/critères*). En ce qui concerne le deuxième nom (ici : *sélection*), il ne subit aucun changement (*\*critère(s) de(s) sélections*). Le déterminant zéro devant *sélection* est stable et ne change pas. Ainsi, nous ne pouvons pas dire *\*critère de la sélection* ou *\*critère des sélections*. Dans notre corpus, il y a deux adjectifs qui vont avec cette collocation : *aléatoire* et *clair*. Nous n'avons pas trouvé d'adjectifs qui peuvent être insérés entre le N1 et le N2. Il y a 6 disciplines dans lesquelles cette collocation apparaît : Science info, Sciences politiques, Sociologie, Psychologie, Économie, Anthropologie.

Au regard des résultats, voici les contraintes observées:

1. La forme typique est au pluriel : *les critères de sélection*. La forme au singulier est présente, mais elle n'est pas très fréquente.
2. Le déterminant qui apparaît le plus souvent devant le N1 est un article défini (*les*). Il arrive que cela soit un article indéfini (*des*) : la différence entre les deux n'est pas vraiment considérable (5 occurrences contre 3).
3. La collocation ne peut subir aucun changement. La préposition *de* est invariable, ainsi que le N2 (ici : *sélection*).

Exemple	N1	déterminant	N2
<b>Critère de sélection</b>	- peu de contraintes	- pas de déterminant pour le N2	- le N2 est en général au singulier -un nom d'idéalité ou un nom d'activité - un nom abstrait

### Forme typique: l'augmentation du nombre

#### Lemme : augmentation du nombre

Les résultats relèvent que la forme typique de la collocation est *l'augmentation du nombre*. Le déterminant qui apparaît avec le N1 au singulier (ici : *augmentation*) est un article défini (*le* – plus précisément sa forme élidée *l'* – 14 occurrences) ou un article indéfini *une* (9 occurrences). Quant à la préposition, c'est un article contracté *du* qui ne change pas. Dans le corpus, nous avons observé l'insertion des adjectifs entre le N1 et le N2. La collocation apparaît dans 6 domaines : Psychologie, Sciences politiques, Géographie, Économie, Histoire, Linguistique, Sociologie, Sciences info.

Quant aux contraintes :

1. La forme typique est au singulier: *l'augmentation du nombre*. Nous n'avons pas remarqué la forme au pluriel pour cette collocation (*\*les augmentations du nombre...*).
2. Le déterminant qui apparaît le plus souvent devant le N1 est un article défini (*le*), plus précisément, sa forme élidée *l'*.
3. Le déterminant avant le N2 ne change pas.

4. Il y a des possibilités d'insérer des adjectifs après le N1(cf. *Une augmentation importante du nombre*) et avant le N1 (*une légère augmentation du nombre*).
5. Il n'y a pas de variations possibles.

Exemple	Remarque	N1	déterminant	N2
<b>Augmentation du nombre</b>	- beaucoup de contraintes	- toujours au singulier	- il ne change pas, la forme contractée (de + le = du)	- le N2 est toujours au singulier

### Forme typique: l'enquête de terrain

#### Lemme : enquête de terrain

Nous pouvons observer que le premier constituant de la collocation **enquête de terrain** varie en nombre (*enquête/enquêtes de terrain*). La forme typique est *l'enquête de terrain*. Le déterminant avant le N1 au singulier est souvent l'article défini *le* ou l'article indéfini *un* (14 contre 12 cooccurrences). Dans notre corpus, nous avons observé également l'utilisation des déterminants possessifs *notre* et *son*. Quant au pluriel, c'est l'article défini *les* qui domine le pion. La préposition *de* n'est pas suivie d'un déterminant. Elle reste la même dans tous les cas. Dans cette collocation, la sélection du N2 est très souvent contrainte. La prép de N2 (ici: *de terrain*) est comprise comme un adjectif. *De terrain* s'applique à la classe des processus scientifiques (*études de terrain, recherche de terrain, travail de terrain*). La détermination de N2 est toujours contrainte, avec un déterminant  $\emptyset$ . Dans le corpus, nous avons vu deux adjectifs qui vont avec cette collocation, à savoir : *journalistique* et *monographique* (cf. *les enquêtes de terrain journalistiques*). La collocation apparaît dans 6 disciplines : Géographie, Anthropologie, Sciences Politiques, Sciences éducation, Sciences info.

Énumérons les contraintes :

1. La forme typique est au singulier: *l'enquête de terrain*. La forme au pluriel est également fréquente : *les enquêtes de terrain*.
2. Le déterminant qui apparaît le plus souvent devant le N1 est un article défini (*le*), pour le pluriel, c'est un article défini *les*.
3. Le N2 n'est pas accompagné de déterminant
4. Il y a des possibilités d'insérer des adjectifs après le N1(cf. *L'enquête monographique de terrain*)

5. Le N2 est toujours au singulier et sa détermination est toujours contrainte.

Exemple	Remarque	N1	déterminant	N2
<b>Enquête de terrain</b>	- <i>De terrain</i> est compris comme un adjectif	- peu de contraintes	- pas de déterminant pour le N2	- le N2 est au singulier - le complément joue un rôle d'adjectif

Nous avons choisi 25 collocations de type N prép N. Nous avons vérifié leur usage dans le logiciel Lexicoscope en utilisant le corpus scientifique. L'emploi de ces collocations ainsi que des contraintes que nous avons dégagées pendant notre étude nous ont permis de pouvoir cibler les difficultés et permettre d'en essayer de présenter une modélisation dans le but de créer un dictionnaire des collocations en ligne.

En général, les mots sont regroupés en classes selon leurs propriétés formelles (morphologiques et syntaxiques) et sémantiques (autrement dit, le type de notion qu'ils véhiculent) (Leeman, 2004).

L'analyse de ces données révèle des résultats intéressants. Néanmoins, il faut ajouter que le nombre des collocations choisies était peu important. En nous basant sur ces exemples, nous proposons un classement :

N1 toujours au singulier (noms déverbaux)	N1 toujours au pluriel	N1 ayant peu de contraintes, c'est-à-dire, il peut apparaître tant au singulier qu'au pluriel	Cas différent
<b>Augmentation</b> du nombre  <b>Partie</b> de l'article  <b>Diffusion</b> d'information	0	Critère de sélection Analyse des données Enquête de terrain Champ d'étude Objectif de recherche Groupe de travail Objet d'étude Période d'observation Point de départ Programme de recherche Recherche en science Réponse à une question Résultat de l'étude Recueil de données Principe d'organisation Point d'appui Outil de communication Objet d'analyse Mode d'action	<b>Processus</b> de construction – la même forme : pour le N1 au singulier et au pluriel

		Ensemble de pratique Valeur de référence	
--	--	---	--

Nous avons classifié nos 25 collocations. Nous avons dégagé des collocations où le N1 est toujours au singulier. C'est le cas de *diffusion d'information*, *partie de l'article* et *augmentation du nombre*. Nous n'avons trouvé aucune collocation qui a le N1 toujours au pluriel. Par contre, il y a beaucoup de collocations où le N1 peut se mettre facilement au pluriel ou au singulier. Dans notre recherche, il y avait un cas particulier de *processus de construction* où le N1 est le même aussi bien au singulier qu'au pluriel.

Dans le tableau suivant, nous avons observé si le N2 peut être au singulier ou au pluriel.

N2 toujours singulier	N2 toujours pluriel	N2 ayant peu de contraintes, il peut être, soit au singulier, soit au pluriel
Objectif de <b>recherche</b> Critère de <b>sélection</b> Augmentation du <b>nombre</b> Enquête de <b>terrain</b> Programme de <b>recherche</b> Groupe de <b>travail</b> Période d' <b>observation</b> Point de <b>départ</b> processus de <b>construction</b> principe d' <b>organisation</b> outil de <b>communication</b> valeur de <b>référence</b>	recueil de <b>données</b> Analyse des <b>données</b>	Champ d'étude recherche en science réponse à une question point d'appui partie de l'article objet d'analyse mode d'action ensemble de pratique diffusion d'information Objectif de recherche

Les résultats démontrent qu'il y a des collocations où le N2 est toujours au singulier, c'est-à-dire qu'il ne se met jamais au pluriel. Parmi les collocations que nous avons analysées, il y en a deux où le N2 est toujours au pluriel, c'est le cas de *recueil de données* et *analyse des données*. Le substantif « données » ne se présente jamais au singulier. Comme dans les cas ci-contre, le complément se met au pluriel quand il comprend plusieurs éléments, un recueil ou une analyse contiennent nécessairement plus qu'une donnée.

Le fait que le groupe nominal ordinaire peut être, soit au singulier, soit au pluriel peut poser certains doutes. Nous pouvons nous demander s'il faut écrire « ensemble de pratique(s) » avec « pratique » (ou un autre nom) au singulier ou au pluriel. Effectivement, il semble qu'après les noms tels que *genre*, *espèce*, *classe*, *forme*, etc., qui sont suivis de la préposition *de*, le nom est parfois au singulier, parfois au pluriel.

Autrement dit, les deux formes sont possibles. Ainsi, il faut prendre en considération le sens ou l'intention que nous voulons donner à la phrase. Comme exemple, nous pouvons dire *un ensemble de pratiques* ( où un ensemble comprend plusieurs pratiques). Si notre intention est d'insister sur une chose en particulier, il faut plutôt utiliser le singulier. Par exemple : « cet ensemble de pratique ». Dans ce contexte, nous insistons sur une pratique en particulier.

Le tableau ci-dessous représente des collocations du type *N1 prep. N2* où les deux substantifs peuvent se mettre au pluriel:

La possibilité d'avoir N1 et N2 PLURIEL
Analyses des données Objets d'études Programmes de recherches Recherches en sciences Réponses à des/aux questions Résultats d'études Recueils de données Points d'appuis Modes d'actions Ensembles de pratiques

Dans la plupart des cas, le N2 peut être considéré comme un complément de nom. Dans les collocations telles que *programme de recherche* et *point d'appui*, les mots *recherche* et *appui* sont appelés *complément du nom*. Nous nous demandons si le complément du nom prend la marque du pluriel, en particulier lorsque le premier nom est lui-même au pluriel? Devons-nous écrire *des programmes de recherche* ou *des programmes de recherches*? Et quand le premier nom est singulier? Quelle graphie privilégier : *un point d'appui* ou *un point d'appuis*? D'où vient à l'aide un dictionnaire des collocations avec des propriétés syntaxiques décrites. Les dictionnaires devraient fournir une forme typique de chaque collocation d'où savoir quelle graphie il faudrait privilégier. En règle générale, il faudrait généralement analyser chaque cas et se poser la question si le complément évoque plutôt l'idée d'un seul élément ou, au contraire, l'idée de plusieurs éléments.

Après avoir suivi les résultats, nous avons observé qu'en règle générale, le complément du nom se met au singulier quand il désigne :



- un seul élément ou la notion : *objet de recherche* (une recherche particulière, on parle bien d'une recherche concrète) ;
- *des enquêtes de terrain* (dans ce cas-là, chaque enquête ne concerne qu'un terrain), le sens générique ;
- une chose non comptable, c'est-à-dire qu'on ne peut dénombrer (qu'elle soit abstraite ou concrète), qui a plutôt le sens général : *valeur de référence*.

Le complément se met au pluriel quand il comprend plusieurs éléments :

- *des analyses de données* (une analyse contient nécessairement plus qu'une donnée)

Répondons maintenant aux questions qui guident notre recherche (cf. partie II, chapitre 2, p. 62,63):

- Nous avons vu qu'il y a des collocations où le N1 peut être toujours au singulier mais il y a des cas où le N1 peut se mettre facilement au pluriel ou rester au singulier (ce sont des collocations ayant « peu de contraintes »)
- Quant au N2, nos résultats étaient partagés. Il y a des collocations où le N2 est toujours singulier (cf. *objectif de recherche*) ou toujours au pluriel (cf. *analyse des données*), mais il y a aussi des cas où le N2 a peu de contraintes et il peut se mettre facilement, soit au singulier, soit au pluriel (cf. *recherche en science/sciences*).
- Le N1 qui peut apparaître, soit au singulier, soit au pluriel est le plus souvent déterminé par l'article défini *le*, *l'* ou *les*. Certes, il y a d'autres déterminants qui apparaissent, mais en règle générale, la forme typique est présente avec un article défini.
- S'il s'agit des adjectifs, la plupart des collocations permettent d'insérer des adjectifs devant ou après le N1 (cf. *diffusion rapide de l'information*) et le N2 (*diffusion de la bonne information, diffusion d'information commerciale*). La seule collocation qui ne subissent pas de changement concerne l'exemple *valeur de référence*.

S'il est question de contraintes concernant la préposition avant le N2, il y a des collocations où il y a un déterminant Ø devant le N2. Voici un tableau :

Déterminant Ø – N2
Critère de sélection
Enquête de terrain
Groupe de travail
Période d'observation
Point de départ
Processus de construction
Recherche en science
Point d'appui
Outil de communication
Mode d'action
Ensemble de pratique
Valeur de référence

Tableau 11 : Déterminant Ø – N2

Quand N1 est un nom du type *sorte*, *espèce* ou *ensemble* (*un ensemble de pratiques*), le N2 peut également changer de nombre, puisque le N1 est une sorte de commentaire métalinguistique sur le choix de N2 : le nombre est déterminé par l'unité ou la pluralité de ce qui est visé sous le vocable à valeur référentielle pratique.

Dans ce chapitre, nous avons essayé d'observer quelles modifications les collocations peuvent subir. Cette étude nous a permis de voir des nuances au niveau sémantique entre différentes structures syntaxiques de la même collocation. Ainsi, nous avons vu quelle est la différence, par exemple, entre *le champ d'étude* et *le champ de l'étude*. Nous pensons qu'il est important de saisir cette nuance vu que le langage scientifique est spécifique et il est important qu'il soit compréhensible pour tout le monde.

## Chapitre 4. Proposition de la modélisation

Le repérage et l'étude sur des propriétés syntaxiques des collocations devraient permettre de déboucher sur une modélisation de ces expressions. Dans cette section, nous nous intéresserons particulièrement à la modélisation des propriétés syntaxiques pour un dictionnaire de collocations.

En se basant sur l'étude des dictionnaires que nous avons vus dans la partie précédente, nous essayons de proposer une modélisation des propriétés syntaxiques pour un dictionnaire de collocations en ligne. Nous soutenons l'idée que la représentation des informations syntaxiques devrait être sous forme de graphiques. En effet, cela rendrait l'accès aux données plus accessible et plus simple. Néanmoins, les propriétés syntaxiques ne sont pas faciles à représenter.

Leur illustration dans le Robert Correcteur est très innovante : rappelons qu'il y a une vue graphique d'un mot employé avec les adjectifs. Plus l'adjectif est gros, plus sa fréquence est importante.

Notre étude est une fenêtre sur la compréhension de la modélisation des collocations dans des dictionnaires. Comprendre ce fait peut aider à dévoiler les restrictions liées à ce problème et ainsi, à réfléchir à créer un dictionnaire des collocations en ligne. Nous avons élaboré notre propre prototype de codage des collocations. Nous y avons codé deux collocations à titre d'exemple *champ d'étude* et *objet d'étude*. Nous avons également essayé de présenter une vue graphique.

Regardons le prototype :

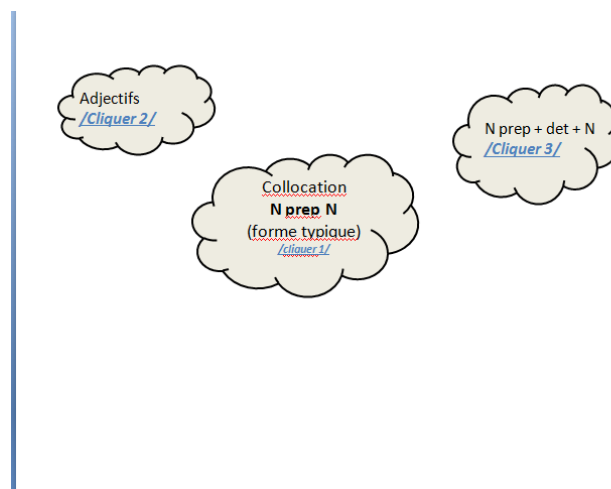


Figure 13 : Prototype : Entrée

Au milieu, se trouve en gras une collocation du type N1 prep N2 qui est sa forme typique. À gauche, nous avons une entrée pour les adjectifs (que l'on peut insérer dans la collocation) et finalement, à droite, celle pour les déterminants qui accompagnent le substantif 2. Pour étudier plus en détail les adjectifs ou les déterminants, il faut cliquer sur l'une des entrées en question. Cela nous permettra de voir la représentation des propriétés syntaxiques des collocations dans plusieurs onglets. Nous sommes pourtant conscients que cela suscite beaucoup de travail. Cependant, nous soutenons l'idée que la clarté et la facilité d'usage sont les plus importantes, surtout pour des apprenants des langues étrangères.

Lorsque nous cliquons sur le nuage 1 (cliquer 1), nous avons un nouvel onglet qui apparaît, le voici :

*/cliquer 1/*

- Lemme :
- Definition:
- Exemple d'usage :

## Disciplines ?

Figure 14 : Prototype 1

Sur cette page, il y a le lemme de la collocation, sa définition et un exemple d'usage. Nous trouvons que l'exemple d'usage est très important du point de vue d'un locuteur non natif. Cela permet de voir une expression dans son contexte naturel, grâce à quoi l'apprenant peut comprendre plus facilement dans quels contextes on peut utiliser telle ou telle expression. Nous nous sommes dit qu'il est important d'ajouter les disciplines dans lesquelles telle ou telle collocation apparaît. Ainsi, nous pouvons observer quelle collocation est typique pour quelle discipline. Les discipline sont classées selon la fréquence.

Après, nous revenons à la vue d'entrée et nous cliquons sur le nuage 2.

## */cliquer 2/*

Légende :

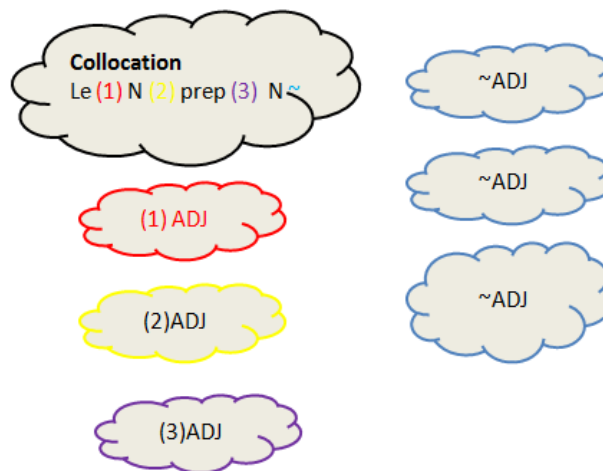


Figure 15 : Prototype 2

Le nuage 2 concerne les adjectifs que l'on peut insérer dans la collocation. Pour que la représentation soit claire, nous avons utilisé plusieurs couleurs. Ainsi, au milieu, nous avons une collocation et des chiffres (entre parenthèses et en couleur) qui indiquent la place d'insertion. Le tilde (~), que nous retrouvons à la fin, désigne des adjectifs postposés au N2.

Dans ce cas-là, nous avons : *Le (adjectif1) nom (adjectif2) prép (adjectif3) N ~(adjectif)*. Nous verrons dans les exemples faits comment le codage présent marche.

Le dernier nuage concerne les déterminants antéposés au N2. Nous pouvons y observer des variations possibles entre les déterminants de N2.

## */cliquer 3/*

- N1 prep det N2

Variations possibles

**Collocation** – signification – exemple d'usage

**Collocation** – signification – exemple d'usage

**Collocation** – signification – exemple d'usage

...

Figure 16 : Prototype 3

Regardons maintenant comment cela marche avec des exemples précis. Prenons un exemple de la liste N1 prép N2 : le **champ d'étude**



Figure 17 : Entrée : champ d'étude

La forme typique de cette collocation est *le champ d'étude*. Lorsque nous cliquons sur le premier nuage, nous obtenons les résultats suivants :

[/cliquer 1/](#)

- **Lemme** : champ d'étude
- **Definition** : C'est un ensemble de connaissances fondées sur diverses disciplines.
- **Exemple d'usage** : C'est le *champ d'étude* de leur statistique de 1848, qui exclut les chantiers et manufactures en banlieue des industriels logés à Paris - et n' envisage pas la situation inverse (Claire Lemerrier, *Revue d'histoire moderne et contemporaine* : 54-3)

Sciences info  
Sociologie  
Histoire  
Géographie  
Sciences education  
Sciences politiques  
Anthropologie  
Psychologie  
Economie

Figure 18 : Entrée : champ d'étude

Nous voyons d'abord le lemme qui est *champ d'étude*. Ensuite, il y a une définition de cette collocation et un exemple d'usage tirés de Lexicoscope. À droite, nous avons les disciplines dans lesquelles la collocation *champ d'étude* apparaît. Nous voyons bien que c'est dans les sciences informatiques où cette collocation est utilisée le plus souvent.

Si nous cliquons sur le nuage 2, nous apercevons quels adjectifs peuvent être insérés dans la structure. Ainsi, nous obtenons ces résultats :

[/cliquer 2/](#)

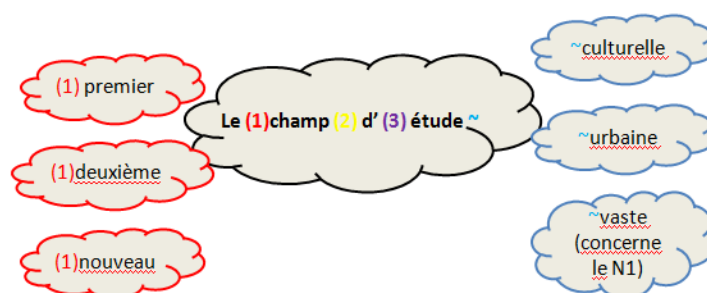


Figure 19 : Prototype 2 : champ d'étude

Il y a 3 adjectifs que nous pouvons insérer avant le N1 : *le **premier** champ d'étude*, *le **deuxième** champ d'étude*, *le **nouveau** champ d'étude*. Il n'y a pas d'adjectifs à mettre après le N1 et avant la préposition *de* ni après la préposition *de* et avant le N2 ((2), (3)). Par contre, le tilde nous montre qu'il y a trois adjectifs que nous pouvons mettre à la fin de la collocation : *le champ d'étude **culturelle***, *le champ d'étude **urbaine***, *le champ d'étude **vaste***.

Quant aux déterminants avant le N2, consultons le nuage 3 :

[/cliquer 3/](#)

- N1 prep det N2

Variations possibles

**Champ de l'étude** — se réfère au domaine d'une étude spécifique — *Dans la partie économétrique, nous restreignons le champ de l'étude aux hommes salariés travaillant dans le secteur privé en 1988 et dont la dernière sélection dans le dispositif de formation a eu lieu dans l'entreprise de 1988 au moins un an avant le début de l'enquête NNN6*

**Champ de cette étude** — la focalisation sur l'étude present, de laquelle on est en train de parler — *Ce phénomène s'est accentué par la suite et les générations suivantes, qui n'entrent pas dans le champ de cette étude, terminent en moyenne leurs études après 20 ans (Chauvel [1998])*

**Champ des études** — le champ concerne plusieurs études et non une seule étude — *L'histoire comparative et le champ des études migratoires*

Figure 20 : Prototype 3 : champ d'étude

Nous voyons les variations possibles de la collocation *champ d'étude*, à savoir : *champ de l'étude, champ de cette étude et champ des études*. Chaque variation est accompagnée d'une définition et d'un exemple d'usage. Nous trouvons cela important, car les variations permettent de se focaliser sur les aspects spécifiques. Tout dépend sur quoi nous voudrions insister et sur quoi porter attention.

Analysons un exemple un peu plus complexe :



Figure 21 : entrée : objet d'étude



## /cliquer 1/

- **Lemme** : *Objet d'étude*
- **Definition** : Un sujet dont on parle, que l'on questionne.
- **Exemple d'usage** : Ce langage va servir non seulement à nommer les éléments d' **un objet d'étude** à créer un lexique, à donner un contenu ( fonction sémantique )

Figure 22 : Prototype 1 : objet d'étude

La première représentation est simple. Il y a un lemme, une définition de la collocation et un exemple d'usage.

Quant aux adjectifs, la situation est plus complexe. Cette collocation peut subir des insertions de différents adjectifs :

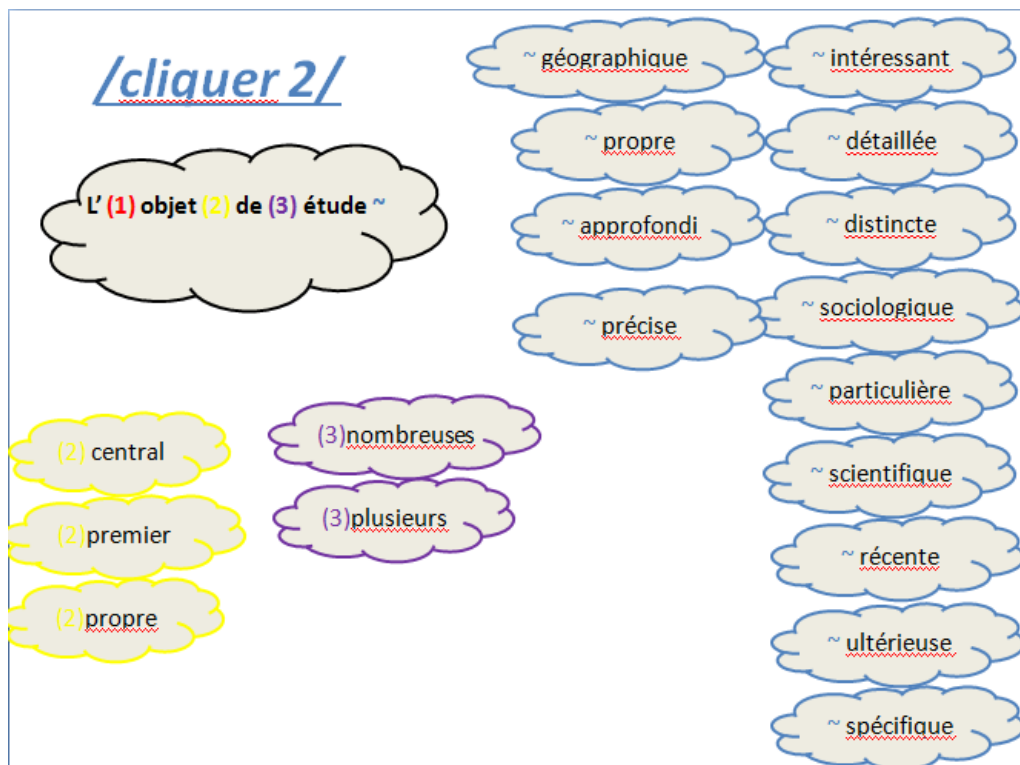


Figure 23 : prototype 2 : objet d'étude

Nous n'avons pas trouvé d'adjectifs que l'on peut insérer avant le N1. Par contre, il y a des adjectifs que l'on peut mettre après le N2 (en jaune), à savoir : *central*, *premier*,

*propre* (l'objet **central** d'étude, l'objet **premier** d'étude, l'objet **propre** d'étude. Il y a des adjectifs (en violet) que l'on peut insérer devant le N2 (l'objet de **nombreuses** études, l'objet de **plusieurs** études). La tilde (~) indique quels adjectifs nous pouvons mettre à la fin de la collocation. Nous pouvons nous demander s'il faut tout mettre ou seulement sélectionner les exemples les plus fréquents. Dans l'exemple ci-dessous, nous avons choisi de tout mettre, mais c'est seulement l'une de propositions du codage.

Le nuage 3 montre les variations possibles des déterminants avant le N2 :

[/cliquer 3/](#)

- N1 prep det N2

Variations possibles

**Objet de cette/notre/ son étude** — la focalisation sur l'étude

présent, de laquelle on est en train de parler — L'existence d'un lien de causalité entre les deux est difficile à démontrer, mais l'historien peut, et c'est l'objet de *cette* étude, s'interroger sur le contexte psychologique dans lequel s'inscrit l'action des décideurs (A. Tachin, Revue historique)

**Objet d'une étude** — la focalisation sur l'étude particulière — Les relations

compliquées entre Rousseau et les représentants ont récemment fait l'objet d'une étude de la part de Richard Whatmore à propos d'un autre texte genevois de Rousseau, les Lettres de la montagne (R. Markovits, Revue historique 652)

**Objet de l'étude** — se réfère au domaine d'une étude spécifique — Les sens

communément admis provenant de l'expérience sociale ne sont pas identiques à ceux ayant pour origine le champ scientifique et le champ académique : ainsi les significations sociales sont-elles construites autant par le champ objet de l'étude que par le champ académique ou celui des médias (J.Ferreira, Les Enjeux de l'information et de la communication Volume 2010)

Figure 24 : Prototypage 3 : Objet d'étude

Il faut bien dire que la représentation de ces données n'est pas évidente. Nous pouvons opter pour la représentation en graphes ou en tableau. Néanmoins, nous proposons un autre codage. En tenant compte que c'est un dictionnaire utilisable en ligne, nous pouvons profiter des possibilités que les nouvelles technologies offrent. Nous sommes conscients que la représentation présente est un peu rude, mais c'est une tentative de la représentation des propriétés syntaxiques des collocations.

## Conclusion

L'étude que nous avons effectuée a permis de faire deux constats : les collocations sont omniprésentes dans chaque langue et leur usage est important. Par ailleurs, les propriétés syntaxiques ne sont pas bien traitées dans les dictionnaires de collocations actuels et la réflexion sur leur représentation n'est pas facile. Il faut ajouter que certains figements syntaxiques sont importants, par exemple *peur bleue*. Nous ne pouvons pas dire *la peur est bleue*, *une peur très bleue* contrairement à un *steak bleu* qui peut être *bleu* même *très bleu*. Pour les collocations N prep N, les associations ne se font pas indifféremment comme nous l'avons démontré. Pour le figement syntaxique, il faut comprendre que parfois nous ne pouvons introduire aucun élément entre les éléments constitutifs de l'expression parce qu'elles n'auront pas le même sens. Ainsi, il faut apprendre ces structures en bloc ou proposer une bonne aide qui pourrait expliquer ces différences. Afin de savoir utiliser correctement différentes collocations (p.ex. *champ d'étude* et *champ de l'étude*), il est indispensable de connaître des nuances au niveau sémantique entre leurs structures. En effet, il existe une petite différence entre ces deux associations de mots. Leur usage dépend de ce que nous voudrions dire et sur quoi nous voudrions focaliser notre attention. La présence du déterminant *l'*, dans ce cas, apporte un autre sens à la collocation. Cela se réfère à toutes les collocations qui peuvent être l'objet de modifications. En effet, l'absence ou l'ajout d'un déterminant peut changer le sens de la collocation. Ces changements peuvent également concerner le N1 ou le N2 qui peuvent être employés au singulier ou au pluriel (*champ d'étude* vs. *champ d'études* vs. *champs d'étude*).

À notre sens, le modèle de notre dictionnaire pourrait fournir une information linguistique détaillée sur les collocations, les relations qu'elles entretiennent entre elles, ainsi que leur usage. Il convient de noter que cela serait utile pour les étudiants, locuteurs non natifs, dans le cadre de la didactique du français de spécialité, mais également pour ceux qui veulent écrire correctement. De plus, notre dictionnaire est représenté de façon claire et accessible à un grand public. Il est donc destiné à chaque utilisateur, que ce soit un enseignant ou un apprenant. La forme de ce dictionnaire est assez simple. En conséquence, elle devient également visible et agréable à utiliser. Grâce à ces avantages, l'accès à des informations concernant les propriétés syntaxiques des collocations est, à notre avis, beaucoup plus facile que dans les dictionnaires dont la forme est compliquée et le codage est trop complexe.

Les résultats de notre travail montrent qu'il existe plus de collocations subissant des modifications que celles restant figées, ce qui empêche sans doute leur utilisation correcte. En effet, le fait que les modifications peuvent être introduites dans certaines collocations ne facilite sûrement pas leur apprentissage. Nous pensons donc que notre recherche ne pourrait pas être exploitée à l'avenir uniquement pour des objets didactiques, mais également pour différents projets scientifiques ayant pour but de créer des dictionnaires. Néanmoins, nous nous rendons compte que notre travail se limite à l'analyse détaillée de 25 collocations, ce qui ne constitue pas un nombre vraiment important. Nous trouvons pourtant que cette étude pourrait influencer de certaines manières la science, car elle traite une question linguistique importante. En effet, les collocations sont depuis toujours présentes dans les langues et notre étude peut donner une idée sur le codage de ces propriétés.

Le sujet de notre mémoire nous paraît être intéressant, car notre recherche traite des collocations tant du point de vue syntaxique que sémantique. De plus, elles sont difficiles à étudier surtout pour les apprenants étrangers du FLE en raison des modifications qu'elles peuvent subir. Notre travail dévoile les difficultés concernant les propriétés syntaxiques des collocations auxquelles les étudiants peuvent être exposés lors de leur utilisation.

Nous nous sommes focalisé avant toute chose sur le champ lexical se référant aux écrits scientifiques. Néanmoins, nous pouvons avancer une hypothèse générale concernant les collocations. L'information syntaxique incluse dans les dictionnaires devrait préciser le type de constructions qui accompagnent le mot (type de complément, etc.), le type de déterminant (qui est souvent complexe à prévoir) et les aspects distributionnels comme la place des adjectifs. Comme nous l'avons démontré dans la partie relative à l'inventaire des dictionnaires, même les dictionnaires spécialisés des collocations ne fournissent pas des informations détaillées à propos des propriétés syntaxiques. Dans les perspectives, nous trouvons que c'est un vide qu'il faut combler. De plus, le langage de spécialité est une langue rigoureuse où les règles sont bien établies. C'est pour cela qu'il faut écrire correctement, pour que tout le monde puisse comprendre ce que nous voulons véhiculer par telle ou telle idée. Dans notre travail, nous voulions rapprocher l'aspect grammatical et sémantique en essayant de repérer les propriétés syntaxiques des collocations. Si nous reconsidérons les problématiques de notre recherche, il s'avère que le sujet n'est pas vraiment bien traité même par des dictionnaires de collocations.

Dans la présente recherche, pour repérer des collocations, nous nous sommes basée sur la liste déjà établie dans le cadre de projet Termith, mais nous avons aussi repéré des constructions en parcourant des articles scientifiques dans le programme Lexicoscope. Une partie de notre travail consistait également à élaborer une modélisation de ces propriétés. Nous avons voulu relier les aspects syntaxiques/grammaticaux aux aspects sémantiques.

Il nous semblait pertinent de prendre en compte l'aspect pragmatique, car cela pourrait servir aux étudiants étrangers prenant connaissance du langage scientifique écrit. En effet, afin de bien maîtriser une langue étrangère, il est nécessaire de savoir quels mots peuvent se lier l'un avec l'autre. Les relations entre les mots ou les termes dont la cooccurrence est élevée dans une langue donnée sont importantes. L'analyse de l'aspect pragmatique des collocations rendra donc leur utilisation plus simple. Ainsi, à côté de la description formelle des collocations, ils seront conscients de leurs emplois réels. Notre dictionnaire fournirait des informations concernant l'usage de chaque collocation ainsi que les disciplines dans lesquelles chaque collocation est utilisée le plus souvent.

Le travail sur ce sujet nous a permis d'explorer les questions comme les rapports entre le genre et la grammaire ainsi que la sémantique. En tant que locutrice non native, nous trouvons important de savoir de se rendre compte de ce problème. En polonais, il n'y a pas d'articles. C'est pourquoi il est intéressant de noter la différence entre *champ d'étude* et *champ de l'étude* ou de savoir que *critère de sélection* possède l'article  $\emptyset$  et ne subit aucun changement (\*critère de la sélection). De plus, l'aspect vraiment intéressant de cette étude était l'exploitation des corpus. Ce travail nous a montré des possibilités et des limites des outils qui servent à extraire automatiquement des informations.

## Bibliographie

- Bally Ch. (1909 [1951]), *Traité de stylistique française*. Genève: Georg et Cie – Paris
- Brunelle E., Bergevin M.F., Fontaine J., Laroche A., Pirronneau M., Papik F., Orsonnens A. (2012), *Guide d'utilisation, Antidote, Posologie*, Société Druides, Canada
- Clear J., (1992), *From Firth Principles, Computational Tools for the Study of Collocation*, Cobuild, Birmingham.
- Firth, J. R. (1957). *Modes of meaning*. *Papers in Linguistics 1934-1951*, Oxford, pp. 190-215.
- Fontenelle, T. (1992), *Collocation acquisition from a corpus or from a dictionary: a comparison*. In *Proceedings I-II. Papers submitted to the 5th EURALEX International Congress on Lexicography in Tampere*, 221-228.
- Grossman F., Tutin A. (2015), *NOUN PREP NOUN collocations in French : the case of scientific lexicon*, In Orlandi A.(éd.), *Defining collocation for lexicographic purposes*, Peter Lang, 269-298.
- Grossman F., Tutin A. (2002), *Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif*, *Revue française de linguistique appliquée*, (Vol. VII).
- Grossmann, F., Tutin, A. (2003), *Les collocations: analyse et traitement.[= Travaux et recherches en linguistique appliquée, Série E: Lexicologie et lexicographie, n° 1]*. Amsterdam: De Werelt.
- Hausmann, F. J., Wiegand, H. E. (1989), *Component Parts and Structures of General Monolingual Dictionaries*, F. J. Hausmann (Ed.), 328-360.
- Hausmann, F. J. (2002), *La transparence et l'obstacle*. *Ela. Études de linguistique appliquée*, (4), 447-454.
- Hausmann F. J., Blumenthal P. (2006), *Collocations, Corpus, Dictionnaires*, *Langue française*, 150.
- Hausmann, F. J. (2007), *Lexicographie française et phraséologie*. In *Collocations, phraséologie, lexicographie*. Aachen: Shaker Verlag, 121-153.
- Heid, U., & Gouws, R. H. (2006). *A model for a multifunctional dictionary of collocations*. In *Atti del XII Congresso Internazionale di Lessicografia: Torino, 6-9 settembre 2006*, 979-988.
- Kraif O., Diwersy S. (2013), *Guide d'utilisation – EmoConc, Applications du Corpus Emolex*.

- Kraif O., Diwersy S. (2012), Le Lexicoscope: un outil pour l'étude de profils combinatoires et l'extraction, *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, volume 2: TALN, 399 - 406,
- Larrivée P. (1997), La structuration conceptuelle du langage (Vol. 86). Peeters Publishers.
- Laurens, M. (1999). La description des collocations et leur traitement dans les dictionnaires. *Disponible en ligne: <http://www.vlrom.be/pdf/994colloc.pdf>*.
- Leeman D. (2004), Les déterminants du nom en français : syntaxe et sémantique, Presses Universitaires de France.
- Legallois D. (2012), La colligation : autre nom de la collocation grammaticale ou autre logique de la relation mutuelle entre syntaxe et sémantique?, *Corpus* (11), 31-54.
- Legallois D., Tutin A. (2013), Vers une extension du domaine de la phraséologie, *Langages* 189
- Lewandowska-Tomaszczyk B., (2012), *Narodowy Korpus Języka Polskiego*. Wydawnictwo Naukowe PWN, Warsaw, 51-58.
- Lewis M., (2000), Teaching Collocation: Further Developments in the Lexical Approach, Hove, England: Language Teaching Publications
- Mel'čuk I. (1998), Collocations and Lexical Functions. In Cowie A. P. (ed.), *Phraseology. Theory, Analysis and Applications*. Oxford : Clarendon Press. 23-53
- Mel'čuk I., (1995), Phrasemes in Language and Phraseology in Linguistics, In Everaert M., van der Linden E.-J., Schenk A., Schreuder R. (éds), *Idioms. Structural and Psychological Perspectives*, Hillsdale, N.J.—Hove, Lawrence Erlbaum Associates, 167-232.
- Mel'čuk I. Clas A., Polguère A.(1995), *Introduction à la lexicologie explicative et combinatoire*. Louvain : Duculot.
- Mel'čuk I. et al. (1984, 1988, 1992, 2000), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I,II, III, IV*. Montréal : Les Presses de l'Université de Montréal.
- Mel'čuk, I. (1998), *Collocations and lexical functions*. In Cowie A. P. (éds.)(2001 [1998]), 23-54.
- Mel'čuk, I. (2011), Tout ce que nous voulions savoir sur les phrasèmes, mais.... *Cahiers de lexicologie, revue internationale de lexicologie et de lexicographie*.
- Meunier F., Granger S. (2008), *Phraseology in Foreign Language Learning and Teaching*, John Benjamin Publishing Company, Amsterdam/Philadelphia

- Nesselhauf, N. (2005) *Collocations in a Learner Corpus*. Amsterdam: John Benjamins Publishing Company.
- Polguère A. (1998), Pour un modèle stratifié de la lexicalisation en génération de texte , *Traitement Automatique des Langues (T.A.L.)*, 39 :2, 1998, 57-76.
- Polguère, Alain, and Igor Mel'auk. "Dérivations sémantiques et collocations dans le DiCo/LAF." *Langue française* 2 (2006): 66-83.
- Selva T., Verlinde S., Binon J. (2003). Le DAFLES, un nouveau dictionnaire électronique pour apprenants du français. *In Proceedings of the Tenth EURALEX International Congress, EURALEX 2002: Copenhagen, Denmark, August 13-17, 2002*, 199-208.
- Sinclair J. (1991), *Corpus, Concordance, Collocations*. Oxford : Oxford University Press.
- Tongini-Bonelli E. (2001), *Corpus Linguistics at Work*, John Benjamin Publishing Company, Amsterdam/Philadelphia.
- Tutin A. (2010), *Sens et combinatoire lexicale : de la langue au discours. Dossier en vue de l'habilitation*
- Tutin A. (2004), Pour une modélisation dynamique des collocations dans les textes. *Actes d'Euralex*. Lorient.
- Tutin A. (2005), Le dictionnaire de collocations est-il indispensable ? *Revue Française de Linguistique Appliquée*.
- Tutin A. (2007), Autour du lexique et de la phraséologie des écrits scientifiques. *Revue française de linguistique appliquée* 12.2, 5-14.
- Tutin A. (2013), Les collocations lexicales: une relation essentiellement binaire définie par la relation prédicat-argument. *Langages* (1), 47-63.
- Tutin A., Kraif O. (2016), Routines sémantico-rhétoriques dans l'écrit scientifique de sciences humaines: l'apport des arbres lexico-syntaxiques récurrents." *Lidil. Revue de linguistique et de didactique des langues* 53, 119-141.
- Vaillant P. (2012). La syntaxe, c'est de la sémantique. *In Documents, Textes, Œuvres (autour de François Rastier)*
- Verlinde S., Binon J., Selva T. (2006), Corpus, collocations et dictionnaires d'apprentissage. *Langue française*, (2), 84-98.
- Wanner L. (1996) (ed.). *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam: J.Benjamins.
- Williams G. (2003), Les collocations et l'École contextualiste. *In Grossmann, F.; Tutin, A. Collocations : analyse et traitement. Travaux et recherches en linguistique appliquée* sÉÉE E 1. 33-44.



## **Dictionnaires :**

*Oxford Collocations Dictionary for Students of English* (2002) (désormais OCDSE) (Crowther *et al.* 2002, réédité et augmenté en 2009, McIntosh. Oxford : Oxford University Press)

*The BBI dictionary of English word combinations* (désormais BBI) de Morton Benson, Evelyn Benson et Robert Ilson (première édition : 1986, réédité en 1993 puis en 1997 et en 2010), Amsterdam/Philadelphia : John Benjamins Publishing Company)

*Le Dictionnaire collocationnel du français général*, de Grobelak, L. (1990) (désormais DCFG). Varsovie, Państwowe Wydawnictwo Naukowe.

*Le Dictionnaire des cooccurrences* (désormais DC), Beauchesne, J. (2001). Montréal, Guérin

*Le Dictionnaire des combinaisons de mots : les synonymes en contexte* (désormais DCM) (sous la direction de Dominique Le Fur, 2007, Paris : Dictionnaire le Robert).

*Le Dictionnaire explicatif et combinatoire*, publié en 4 volumes entre 1984 et 1999 (désormais DEC)

Bańko M. (2006), *Dictionnaire SŁOWNIK DOBREGO STYLU, CZYLI WYRAZY, KTÓRE SIĘ LUBIĄ*

*Larousse de poche 2012, Les éditions Larousse*

## **Dictionnaires en ligne :**

**TLF** : Trésor de la Langue française : <http://atilf.atilf.fr/tlf.htm>

**CNRTL** : Centre National de Ressources Textuelles et Lexicales : <http://www.cnrtl.fr/>

**Larousse** en ligne : [www.larousse.fr](http://www.larousse.fr) : encyclopédie et dictionnaires gratuits en ligne

**ATILF**, Analyse et Traitement Informatique de la Langue Française – [www.atilf.fr](http://www.atilf.fr)

## **Sigles et abréviations utilisés**

N – nom

ADJ – adjectif

N prép N – nom préposition nom

V – Verbe

ADV – adverbe

FL – fonctions lexicales d'Igor Mel'čuk

## Table des illustrations

Tableau 1. Exemple <bloc de mots>.....	10
Tableau 2. Tableau récapitulatif .....	13
Tableau 3. Schéma de Bally.....	15
Tableau 4. Schéma de Galisson.....	15
Tableau 5. Schéma de Hausmann.....	16
Tableau 6. schéma de Mel'čuk.....	18
Tableau 7. Différents types de collocations.....	25
Tableau 8. Classement des collocations N1.....	75
Tableau 9. Classement N2.....	76
Tableau 10. N1 N2 Pluriel .....	77
Tableau 11. Déterminant 0 – N2.....	79
Figure 1. Article de <i>decision</i> dans l'OCDSE.....	29
Figure 2. Article de <i>decision</i> dans le BBI.....	31
Figure 3. Article de <i>décision</i> dans le DCFG.....	33
Figure 4. Article de <i>feuille</i> dans le DC.....	35
Figure 5. Article de <i>décision</i> dans le DCM.....	36
Figure 6. Article de <i>décision</i> dans le DCM : la suite.....	37
Figure 7. Dictionnaire des cooccurrences d'Antidote : <i>décision</i> .....	42
Figure 8. Dictionnaire des cooccurrences d'Antidote, la suite : <i>décision</i> .....	43
Figure 9. Dictionnaire des combinaisons de Robert Correcteur : <i>décision</i> .....	46
Figure 10. Dictionnaire des combinaisons de Robert Correcteur : la suite : <i>décision</i> .....	46
Figure 11. Article de <i>aktywność</i> ( <i>activité</i> ) dans le Słownik Dobrego Stylu.....	49
Figure 12. Arbre dans le programme Lexicoscope : analyse syntaxique.....	58
Figure 13. Prototype : entrée .....	80
Figure 14. Prototype 1.....	81
Figure 15. Prototype 2.....	82
Figure 16. Prototype 3.....	82
Figure 17. Entrée : champ d'étude.....	83
Figure 18. Prototype 1 : champ d'étude.....	83
Figure 19. Prototype 2 : champ d'étude.....	84
Figure 20. Prototype 3 : champ d'étude.....	85
Figure 21. Entrée : objet d'étude.....	85
Figure 22. Prototype 1 : objet d'étude.....	86
Figure 23. Prototype 2 : objet d'étude.....	86
Figure 24. Prototype 3 : objet d'étude.....	87

## Table des annexes<sup>5</sup>

Annexe 1 Les requêtes utilisées pendant notre analyse : .....	98
Annexe 2 Les grilles d'analyses :.....	100
Annexe 3 Tableau récapitulatif .....	156

---

<sup>5</sup> Les annexes doivent être annoncées dans le texte principal dans le corps du texte ou en note de bas de page. On indiquera le n° de l'annexe et la page à laquelle elle se situe. On peut ici détailler ou illustrer des informations qui n'ont pas pu être développées dans le texte mais qui méritent de l'être. Les annexes sont numérotées et titrées. On évitera donc de faire figurer plusieurs annexes sur une même page. Si les annexes sont très nombreuses, il est préférable de les rassembler dans un second volume. Pour enlever cette note de bas de page, supprimer l'appel de note ci-dessus.

## Annexe 1

### Les requêtes utilisées pendant notre analyse :

#### 1. Sans contraintes :

<l=objectif,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<l=recherche,c=NOUN,#3>::(  
PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)

*L = lemme (dans ce cas-là, soit objectif (sg) soit objectifs (pl))*

*C=catégorie*

*(PREPOBJ), (...) – les relations de dépendances entre les éléments*

#### 2. N1 pluriel

<w=objectifs,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<l=recherche,c=NOUN,#3>  
::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)

*W = forme graphique (ici : le premier nom au pluriel)*

#### 3. N2 pluriel

<w=objectif,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<w=recherches,c=NOUN,#3>  
>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)

*Le premier nom au singulier, le deuxième nom au pluriel*

#### 4. DET + N1 singulier

<w=objectif,#1>&&<l=de,c=PREP,#2>&&<l=recherche,c=NOUN,#3>&&<c=DET,#4>::(PREPOBJ,3,2) (\*,1,4)(\* ,1,3)

#### 5. DET + N1 pluriel

<w=objectifs,#1>&&<l=de,c=PREP,#2>&&<l=recherche,c=NOUN,#3>&&<c=DET,#4>::(PREPOBJ,3,2) (\*,1,4)(\* ,1,3)

## 6. N2 avec déterminant

<l=objectif,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<l=recherche,c=NOUN,#3>&  
&<c=DET,#4>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)(DETERM,3,4)

## 7. N1 pluriel N2 pluriel

<w=objectifs,c=NOUN,#1>&&<l=de,c=PREP,#2>&&<w=recherches,c=NOUN,#  
3>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)

## 8. (ADJ1) N1 (ADJ1) de N2

<l=objectif,c=NOUN,#1>&&<l=.\*,c=ADJ,#2>&&<l=de,c=PREP,#3>&&<l=le,c=  
DET,#4>&&<l=recherche,c=NOUN,#5>::(DETERM,5,4)

## 9. N1 de ADJ2 de N2

<l=objectif,c=NOUN,#1>&&<c=ADJ,#4>&&<l=de,c=PREP,#2>&&<l=recherche  
,c=NOUN,#3>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)(.\*,2,4)

## 10. N1 de N2 ADJ

<l=objectif,c=NOUN,#1>&&<c=ADJ,#4>&&<l=de,c=PREP,#2>&&<l=recherche  
,c=NOUN,#3>::(PREPOBJ,3,2) (U3\_DE\_NMOD,1,3)(.\*,3,4)

**Annexe 2**  
**Les grilles d'analyses :**

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Objectif de recherche	30	4 Psychologie Sciences info Anthropologie Sociologie	Et si la pensée cassirérienne est bien profondément génétique , sa reprise dans un cadre scientifique contemporain , de type dynamiciste et constructiviste , constitue un <b>objectif de recherche</b> bien actuel.	author => Jean Lassègue title => L'Homme 192
N1 pluriel	Objectifs de recherche	7	2 Psychologie Sciences info	Ces derniers font état d' un besoin d' information tangible - à savoir une envie de chercher de l' information et du savoir en général - et d' <b>objectifs de recherche</b> relativement clairs , choisis et évolutifs .	author => Karine Aillerie title => Documentaliste-Sciences de l'Information
N2 pluriel	Objectif de(s) recherches	0	0	-	-
DET + N1 singulier	DET objectif de(s) recherche(s)	27 Le : 22 Notre : 1 Son : 2 Un : 2	4 Psychologie Anthropologie Sciences info Sociologie	L' <b>objectif de</b> cette <b>recherche</b> n' est pas d' examiner l' influence des dimensions de la satisfaction sur les comportements de citoyenneté organisationnelle .	author => Pascal Paillé title => Le travail humain
DET + N1 pluriel	DET objectifs de recherche	5 Les : 3 Ses : 1 Des : 1	2 Psychologie Sciences Info	<b>Les objectifs de la recherche</b> étaient présentés dans le cadre de ces groupes de formation .	author => A. Somat title => Le travail

					humain
N2 avec déterminant	Objectif(s) de DET recherche(s)	24 13 :Objectif de <i>cette</i> recherche 10: Objectif de <i>la</i> recherche 1 : Objectif de <i>sa</i> recherche	2  Psychologie Anthropologie	Néanmoins , et c' est bien l'objectif de la recherche, les résultats suggèrent que la réussite des jeunes enfants aux épreuves dites de rotation mentale dépend fortement des caractéristiques des objets et des transformations opérées .	-
N1 pluriel N2 pluriel	Objectifs de(s) recherches	0	0	-	-
N1 +ADJ1 de N2	(ADJ) Objectif (ADJ) de recherche	4 Objectif <i>réel</i> de la recherche Objectif <i>général</i> de la recherche	1  Psychologie	Après avoir rempli le questionnaire , toutes les participantes étaient longuement informées de l' <b>objectif réel de la recherche</b> et pouvaient , si elles le souhaitaient , interroger l' expérimentatrice .	-
N1 de Adj2 N2	Objectif de ADJ recherche	0	0	-	-
N1 de N2 Adj2	Objectif de recherche ADJ	0	0	-	-
	<b>N prép N</b>	<b>Nombre total d'occurrences</b>	<b>Dispersion</b>	<b>Exemple</b>	<b>Métadonnées</b>
Sans contraintes	Critère de sélection	15	6  Science info Sciences politiques Sociologie Psychologie Économie anthropologie	L'un de trois <b>critères de sélection</b> des participants à notre enquête supposait l' accès à l' internet à domicile [...]	author => Karine Aillerie  title => Documentalist e-Sciences de l'Information
N1 pluriel	Critères de sélection	14	6  Science info	Ce résultat est cohérent avec les <b>critères de sélection</b> des adjectifs utilisés pour contrôler la	-



			Sciences politiques Sociologie Psychologie Économie anthropologie	validité de la procédure d' induction émotionnelle .	
N2 pluriel	Critère de(s) sélection	0	0	-	-
DET + N1 singulier	DET critère de sélection	1 : le	1 sociologie	Le critère de sélection était la désignation par les médecins du travail du caractère « précarisé » de la situation des différentes personnes .	author => Valentine Hélardot  title => Cahiers internationaux de sociologie
DET + N1 pluriel	DET critères de sélection	9  les : 5 des : 3 plusieurs : 1	6  Sciences politiques Psychologie Sciences info Sociologie Économie Anthropologie	En pratique <b>les critères de sélection</b> aient d' un collège à l' autre mais incluent le plus souvent les résultats obtenus à l' école élémentaire [...]	author => Gabrielle Fack  title => Actes de la recherche en sciences sociales 180
N2 avec déterminant	Critère de DET sélection	0	0	-	-
N1 pluriel N2 pluriel	Critères de(s) sélections	0	0	-	-
N1 +ADJ1 de N2	Critère ADJ de sélection	0	0	-	-
N1 de Adj2 N2	Critère de ADJ sélection	0	0	-	-
N1 de N2 Adj2	Critère de sélection ADJ	2 Critère de sélection <b>partisane</b> Critère de sélection <b>aléatoires</b>	1  Sciences politiques	Si , traditionnellement , le choix de candidats d' origine populaire a pu distinguer les partis ouvriers , la prise en compte récente , dans un pays comme la France , de l' origine ethnique des candidats ou de la couleur	author => Céline Braconnier  title => Revue française de science politique Vol.

				de leur peau dans le but assumé d'assurer une meilleure représentation du pays et de sa « diversité » dans l'arène politique peut également être lue comme un indicateur d'ethnicisation des <b>critères de sélection partisane.</b>	60
--	--	--	--	--	----

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Analyse(s) de (DET) donnée(s)	41	6 Psychologie Sciences éducation Sciences Info Sciences Politiques Sociologie Géographie	L'approche quantitative s'appuie sur des données d'enquête et utilise <b>l'analyse de données</b> pour , d'une part , identifier les caractéristiques des internautes ayant des amis virtuels , notamment en termes de vie sociale réelle , et , d'autre part , caractériser ces amitiés virtuelles .	author => Virginie Lethiais  Title => Réseaux 164
N1 pluriel	Analyses de donnée(s)	5	3 Psychologie Sciences éducation Sciences politiques	Deux analyses des données sur les comportements pro- puis antisociaux ont validé deux échelles comprenant chacune une structure factorielle en 4 dimensions .	author => V. Leoni  title => Le travail humain
DET + N1 singulier	DET analyse de(s) donnée(s)	24 23 : le 1: un	6 Psychologie Sciences politiques Sciences info Géographie Sciences éducation Sociologie	Il faut donc reprendre <b>l'analyse de données</b> commerciales dans des créneaux technologiques de niveau plus élevé pour voir dans quelle mesure il convient de fortement nuancer les conclusions qui pouvaient être tirées de l'analyse de ces secteurs intensifs en travail .	author => Claude Grasland  title => L'Espace géographique Vol. 39
DET + N1 pluriel	DET analyses de(s) donnée(s)	1 1: leurs	1 Psychologie Sciences Education Sciences politiques	Par <b>leurs analyses secondaires des données</b> Eurobaromètre mais aussi par leur collaboration directe au programme , les chercheurs universitaires ont	author => Philippe Aldrin  title => Politix 89

				largement contribué à lui assurer ce quasi-monopole d'expertise sur l'analyse empirique des attitudes européennes .	
N2 avec déterminant	Analyse(s) de DET donnée(s)	37 35 : des 1 : ces 1 : diverses	6  Psychologie Sciences education Sciences politiques Géographie Sociologie Sciences info	<b>L'analyse de ces données</b> indique une interaction significative entre la congruence frontière syllabique (...)	-
N2 pluriel	Analyse de(s) données	33	6  Psychologie Sciences education Sciences info Sciences politiques Sociologie Géographie	L'analyse des données s'est ensuite focalisée sur le lien entre le fait de percevoir un aspect bloquant / stimulant et la réalisation ou non d'un diagnostic de stress .	author => I. Hansez  title => Le travail humain
N1 pluriel N2 pluriel	Analyses de(s) données	5	3	Deux types d'analyses des données ont été réalisées : la première relève d'une analyse terminologique et thématique , afin de dégager les thématiques et termes récurrents pour caractériser les perceptions et vécus dans les espaces de travail .	author => V. Moulin  title => Le travail humain
N1 +ADJ1 de N2	Analyse ADJ de(s) donnée(s)	9 Analyse <b>qualitative</b> ... Analyse <b>secondaires</b> Analyse <b>statistique</b> Analyse <b>inductive</b> Analyse <b>structurale</b>	4  Sciences éducation Psychologie Sciences politiques Sciences info	L'étude tente d'une part de caractériser ces images identitaires en utilisant une méthodologie <b>d'analyse inductive des données</b> brutes.	-
N1 de Adj2 N2	Analyse de(s) ADJ donnée(s)	0	0	-	-
N1 de N2 Adj2	Analyse de(s) donnée(s) ADJ	11  Analyse des données <b>brutes</b> Analyse des	5  Sciences education Géographie Psychologie Sciences info	Il faut donc reprendre <b>l'analyse des données commerciales</b> dans des créneaux technologiques de niveau plus élevé pour voir dans quelle mesure il	author => Claude Grasland  title => L'Espace géographique

		données <b>agrégées</b> Analyse des données <b>temporelles</b> Analyse des données <b>disponibles</b> Analyse des données <b>contextuelles</b> Analyse des données <b>commerciales</b> Analyse des données <b>bibliographiques</b>	Sciences politiques	convient de fortement nuancer les conclusions qui pouvaient être tirées de l'analyse de ces secteurs intensifs en travail.	Vol. 39
--	--	---	---------------------	--	---------

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Augmentation du nombre	27	8 Psychologie Sciences politiques Géographie Économie Histoire Linguistique Sociologie Sciences info	Parallèlement à l' <b>augmentation du nombre</b> d'espèces végétales, n note une croissance nette de la population d'oiseaux.	-
N1 pluriel	Augmentations du nombre	0	0	-	-
DET + N1 singulier	DET Augmentation du nombre	23 Le : 14 Une : 9	7 Psychologie Géographie Sciences Politiques Histoire Economie Linguistique Sociologie	En revanche , elle montre que l' <b>augmentation du nombre</b> de mises à jour ( 2 versus 4 ) n' influence pas de manière significative les latences au rappel .	-
DET + N1 pluriel	DET augmentations du nombre	0	0	-	-

N2 avec déterminant	Augmentation de DET nombre	1 De <i>leur</i> nombre : 1	1 Sciences politiques	Ces violences sont donc plus souvent enregistrées et statistiquement plus présentes , sans que cela implique une <b>augmentation effective de leur nombre</b> dans la réalité sociale .	author => Pierre Favre  title => Revue française de science politique Vol. 59
N2 pluriel	Augmentation de(s) nombres	0	0	-	-
N1 pluriel N2 pluriel	Augmentations de(s) nombres	0	0	-	-
N1 +ADJ1 de N2	Augmentation ADJ du nombre	6 Augmentation <b>modérée</b> du nombre Augmentation <b>significative</b> du nombre Augmentation <b>importante...</b> Augmentation <b>effective</b> Augmentation <b>générale</b> Une <b>légère</b> augmentation du nombre	4 Psychologie Sciences politiques Histoire Sociologie	Notons qu'une <b>augmentation significative du nombre</b> de sourcils concaves produits apparaît entre 11 ans ( 4/11 ) et l' âge adulte ( 11/13 ) ( p - , 05 ) .	-
N1 de Adj2 N2	Augmentation du ADJ nombre	0	0	-	-
N1 de N2 Adj2	Augmentation du nombre ADJ	1 Augmentation du nombre <i>total</i> : 1	1 Géographie	Si l'on observe la variation des entrées en Grèce pour les ressortissants de différents pays balkaniques ( fig. 6 ) , on constate certes <b>l'augmentation du nombre total</b> d'entrées , qui accrédi- te bien la mise en place d' un comportement migratoire , mais surtout il apparaît une variabilité surprenante pour certaines	author => Pierre Sintès  title => L'Espace géographique

				de ces courbes .	
	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Enquête de terrain	38	5 Géographie Anthropologie Sciences Politiques Sciences éducation Sciences info	Un an après, <b>l'enquête de terrain</b> contribue a faire revivre la campagne.	author => Marie Cartier  title => Politix 91
N1 pluriel	Enquêtes de terrain	19	5 Sciences education Antropologie Science sinfo Géographie Sciences politiques	La répartition entre établissements publics et privés a été possible , avec une faible marge d'erreurs , grâce aux <b>enquêtes de terrain</b> , à l' utilisation du dictionnaire encyclopédique Madras ( 1996 ) et à la série d' articles publiés par le journal local France - Antilles du 11 août au 24 septembre 1999 .	19
DET + N1 singulier	DET Enquête de terrain	18 Une : 10 L' : 8	4 Sciences politiques Géographie Sciences info Anthropologie	À l'issue de <b>l' enquête de terrain</b> , plusieurs informations de natures différentes sont produites , dont le support du zonage sur lequel sont reportées toutes les structures spatiales identifiées par les nomades .	author => Khadija Ait-Alhayane  title => L'Espace géographique Vol. 39
DET + N1 pluriel	DET Enquêtes de terrain	11 Les : 6 Des : 2 Nos : 1 Plusieurs : 1 Ses : 1	4 Sciences education Anthropologie Géographie Sciences politiques	La question ici posée par Schaeffner illustre , parmi d'autres exemples , une approche sociologique des instruments de musique que <b>plusieurs enquêtes de terrain</b> en Afrique ont	author => Brice Gérard  title => L'Homme 191

				rendue concevable , en permettant une observation directe et riche d' instruments [...]	
N2 avec déterminant	Enquête de DET terrain	0	0	-	-
N2 pluriel	Enquête de(s) terrains	0	0	-	-
N1 pluriel N2 pluriel	Enquêtes de(s) terrains	0	0	-	-
N1 +ADJ1 de N2	Enquête ADJ de terrain	2	2	Le Play et les Playsiens cherchent à conjuguer une approche voisine du positivisme avec un empirisme fondé sur la prééminence de l'observaion concrète et de l'enquête <i>monographique</i> de terrain.	author => Lionel Charles title => Annales de géographie
N1 de Adj2 N2	Enquête de ADJ terrain	0	0	-	-
N1 de N2 Adj2	Enquête de terrain ADJ	0	0	-	-

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Champ d'étude	20	6  Sciences education Géographie	C'est le <b>champ d'étude</b> de leur statistique de 1848 , qui exclut les chantiers et manufactures en	author => Claire Lemerrier title => Revue

			Économie Sociologie Sciences info Histoire	banlieue des industriels logés à Paris - et n' envisage pas la situation inverse [ 49 ] .	d'histoire moderne et contemporaine 54-3
N1 pluriel	Champs d'étude	2	2  Sciences education Sciences info	Même si la parole des sujets reste notre point d' ancrage , leurs représentations sont recueillies selon des modalités distinctes dans les deux <b>champs d'étude</b> .	-
DET + N1 singulier	DET Champ d'étude	15 L=Le : 11 Ce : 2 Notre : 1 Un : 1	6 Géographie Economie Sciences education Sociologie Histoire Sciences info	Certes , l' intégration à l' american way of life et au melting pot ne fait plus partie du décor de la société américaine depuis longtemps , mais <b>le champ d'étude</b> reste le même.	author => Gilles Sénécal  title => Annales de géographie
DET + N1 pluriel	DET Champs d'étude	1 1 : les	1 Sciences education	Même si la parole des sujets reste notre point d'ancrage, leurs représentations sont recueillies selon des modalités distinctes dans <b>les deux champs d'étude</b> .	-



N2 avec déterminant	Champ de DET étude	5 Champ de l'étude : 2 Champ de cette étude : 1 Champ des études : 2	2 Economie géographie	Ce phénomène s' est accentué par la suite et les générations suivantes , qui n' entrent pas dans le <b>champ de cette étude</b> , terminent en moyenne leurs études après 20 ans ( Chauvel [ 1998 ] ) .	-
N2 pluriel	Champ de(s) études	2 Champ des études : 2	1 Géographie	Dans les années quatre - vingt , dans le <b>champ des études</b> urbaines une série de travaux dégagent un modèle d' évolution des villes .	author => Sylvie Fol title => Annales de géographie 674
N1 pluriel N2 pluriel	Champs de(s) études	0	0	-	-
N1 +ADJ1 de N2	Champ ADJ d'étude	7 Premier champ d'étude Deuxième champ d'étude ... Nouveau champ d'étude...	3 Sciences education Géographie Sciences info	Ce <b>nouveau champ d'étude</b> dans le domaine de la conservation souffre pourtant d' un manque de définition et ses enjeux n' ont pas été encore clairement synthétisés .	author => Laurent Godet title => L'Espace géographique Tome 39
N1 de Adj2 N2	Champd de ADJ étude	0	0	-	-
N1 de N2 Adj2	Champ d'étude ADJ	3 Champ des études <b>culturelles</b> Champ d'étude suffisamment <b>vaste</b> Champ des études <b>urbaines</b>	2	L' école de Los Angeles se situe sur un autre versant de la géographie , soit du côté des approches qualitatives , littéraires pour une bonne part , et dans le <b>champ des études culturelles</b> d' orientation postmoderne .	author => Gilles Sénécal title => Annales de géographie

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Groupe de travail	80	8 Sciences politiques Sociologie Sciences info Psychologie Sciences education Anthropologie Géographie Histoire	Pour la représentation des scores , <b>le groupe de travail</b> fait le choix d' un code couleur [ 52 ] .	author => Pauline Ravinet  title => Revue française de science politique Vol. 61
N1 pluriel	Groupes de travail	29	7 Sociologie Sciences politiques Anthropologie Psychologie Sciences education Sciences info Histoire	Dans un premier temps , deux <b>groupes de travail</b> sont créés et vont fonctionner jusqu' en 1989 , avant que de nouveaux groupes et de nouvelles initiatives ne prennent le relais ..	author => Karine Aillerie  title => Documentaliste-Sciences de l'Information
DET + N1 singulier	DET Groupe de travail	35 Le : 19 Un : 13 Ce : 2 Son : 1	5 Sciences politiques Sciences info Psychologie Sciences education Sociologie	Le noyau du réseau qui structure <b>le groupe de travail</b> est composé de chercheurs et de professionnels qui se sont rencontrés dans le cadre d' un projet de santé communautaire , dans le quartier Franc - Moisin de Saint-Denis , autour d' un médecin , Antoine Lazarus , et d' un sociologue , Michel Joubert .	author => Didier Fassin  title => Politix
DET + N1 pluriel	DET Groupes de travail	8 des : 5 ces : 1 les : 1 au centre de <b>maints</b> groupes de travail : 1	4 Sociologie Sciences politiques Anthropologie Histoire	Celles-ci seraient préparées tout au long de l' année dans <b>des groupes de travail</b> et des séminaires et les conclusions seraient présentées par deux rapporteurs au début de chaque session .	author => Charles Macdonald  title => Ethnologie française Vol. 38
N2 avec déterminant	Groupe(s) de DET travail	0	0	-	-

N2 pluriel	Groupe de(s) travaux	0	0	-	-
N1 pluriel N2 pluriel	Groupes de(s) travaux	0	0	-	-
N1 +ADJ1 de N2	Groupe(s) ADJ de travail	3 Cinquième groupe Groupe réaliste Petits groupes de travail	2 Histoire Science politiques	La délibération qu' ils organisent s' incarne dans l' animation de <b>petits groupes de travail</b> composés d' habitants et de professionnels ( enseignants , travailleurs sociaux , policiers , etc. ) .	author => Marion Carrel  title => Politix
N1 de Adj2 N2	Groupe de ADJ travail	0	0	-	-
N1 de N2 Adj2	Groupe de travail ADJ	7  Groupe de travail <b>ad hoc</b> Groupe de travail <b>commun</b> Groupe de travail <b>interdépartemental</b> Groupe de travail <b>inexperimenté</b> Groupe de travail <b>interministériel</b>	3  Sciences politiques Géographie Sciences info	Elle souligne avec insistance les inconvénients de sa nouvelle position statutaire : non seulement celle-ci suppose un travail de gestionnaire au détriment du travail d' enquête , mais en outre elle est tenue de diriger <b>un groupe de travail inexpérimenté</b> et d' assumer un leadership pour lequel elle ne se sent guère préparée .	author => Guillaume le Saulnier  title => Réseaux 165

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Objet d'étude	71	10  Linguistique Sciences info Sociologie Histoire Géographie Sciences education Sciences politiques Anthropologie	Ce langage va servir non seulement à nommer les éléments d' <b>un objet d'étude</b> à créer un lexique , à donner un contenu ( fonction sémantique ) ( 4 ) , si possible extrêmement précis , rigoureusement défini , mais aussi à nous	-

			Psychologie Economie	renseigner sur les rapports entre les éléments , c'est-à-dire sur la structure de l'objet étudié ( fonction syntaxique ) .	
N1 pluriel	Objets d'étude	8	4  Sociologie Sciences politiques Géographie Sciences info	Un discours scientifique ne peut pas fonctionner comme un bloc : on doit pouvoir l' amender , l' ajuster sur la base de débats dans lesquels sont mobilisés des arguments faisant une place aux <b>objets d'étude.</b>	author => Michel Grossetti  title => L'Année sociologique Vol. 56
DET + N1 singulier	DET Objet d'étude	53  Le (l') : 33 Un : 14 Notre : 5 Leur : 1	10  Linguistique Histoire Sciences education Géographie Sociologie Anthropologie Economie Psychologie Sciences politiques	L' analyse sémio -discursive qui nous intéresse ici permet d' aborder <b>notre objet d'étude</b> talk-show en tant que « produit fini » , en considérant son « espace interne » d' organisation du discours .	author => Joëlle Desterbecq  title => Etudes de communication
DET + N1 pluriel	DET Objets d'étude(s)	3  différents : 1 les : 1 des : 1	2  Sociologie Sciences politiques	Leur rôle se limite à des suggestions sur <b>les objets d'étude</b> en amont , à des inflexions en cours de recherche et parfois à porter les résultats à la connaissance du CNLE , instance à laquelle il reviendra d' en débattre et d' en faire la matière d' une dispute politique	author => Yves Lochard  title => Politix

N2 avec déterminant	Objet(s) de DET étude(s)	18 Objet de <b>cette</b> étude : 7 Objet d' <b>une</b> étude : 3 Objet de l'étude : 2 Objet <b>des</b> (de+les) études : 2 Objet d' <b>aucune</b> étude : 1 ? Objet de <b>notre</b> étude : 1 Objet de <b>plusieurs</b> études : 1 Objet de <b>son</b> étude : 1	8 Histoire Linguistique Sciences education Sciences info Sociologie Anthropologie Economie Géographie	L'existence d'un lien de causalité entre les deux est difficile à démontrer, mais l'historien peut, et c'est l' <b>objet de cette étude</b> , s'interroger sur le contexte psychologique dans lequel s'inscrit l'action des décideurs.	author => Agnès Tachin title => Revue historique
N2 pluriel	Objet de(s) études	15	8 Linguistique Géographie Histoire Psychologie Sciences Politiques Sociologie Economie Sciences Info	Nous manipulons le lien catégoriel parce que contrairement au lien sur la couleur il a déjà fait l' <b>objet d'études</b> expérimentales.	-
N1 pluriel N2 pluriel	Objets de(s) études	3	1 Sociologie	-Les trois dimensions - de masse, de durée, de généralité - permettent de définir plus précisément les objets sociaux et constituent un bon outil de travail, tant pour construire des <b>objets d'études</b> , que pour exercer la réflexivité qu'implique la recherche en sociologie.	author => Michel Grossetti title => L'Année sociologique Vol. 56
N1 +ADJ1 de N2	Objet(s) ADJ de(s) étude(s)	4 Objet <b>central</b> de l'étude Objet <b>premier</b> de cette étude Objet <b>propre</b>	5 Linguistique Anthropologie Géographie Sciences education Sciences info	<b>L'objet premier de cette étude</b> est d'analyser la mise en place du méta-discours sur le monologue intérieur chez Dujardin en interrogeant ses silences et ses contradictions.	-

		de cette étude			
N1 de Adj2 N2	Objets de ADJ étude	8  Objet de (très) nombreuses études Objet de plusieurs études	8  Géographie Linguistique Histoire Economie Psychologie Sciences politiques Sciences info Sciences education	Il a fait <b>l'objet de nombreuses études</b> à la fin des années 70 et au début des années 80.	-
N1 de N2 Adj2	Objet(s) d'étude(s) ADJ	14  Objet d'étude <b>intéressant</b> Objet d'une étude (plus) <b>détaillée</b> Objet de deux études <b>distinctes</b> Objet des études <b>sociologiques</b> Objet d'une étude <b>particulière</b> Objet de (nombreuses) études <b>scientifiques</b> Objet d'études <b>expérimentales</b> Objet d'aucune étude <b>récente</b> Objet d'études <b>ultérieures</b> Objet d'étude <b>spécifique</b> Objet d'étude <b>géographique</b> Objet d'étude <b>propre</b> Objet d'études <b>approfondies</b> Objet de (plusieurs) études <b>précises</b>	8  Géographie Linguistique Histoire Economie Psychologie Sciences politiques Sciences info Sciences education	Nous manipulons le lien catégoriel parce que contrairement au lien sur la couleur il a déjà fait <b>l'objet d'études expérimentales</b> .	-

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Période d'observation	34	14 Sciences info Sciences education Economie Psychologie	Après une <b>période d'observation</b> de trois semaines , deux assistants de recherche , en stage au sein de l'entreprise , ont distribué différents questionnaires aux salariés .	author => N. Gillet  title => Le travail humain
N1 pluriel	Périodes d'observation	5	1 Sciences education	Complétant ces résultats , il apparaît que , sur les 11 élèves qui n' ont pas été présents lors des trois <b>périodes d'observation</b> , 4 ont conservé la même typologie , 6 ont été classés dans des groupes différents et seul un élève n' a été présent qu' une fois .	author => Pierre Roque  title => Revue française de pédagogie
DET + N1 singulier	DET Période d'observation	28 La : 14 Notre : 9 une : 4 sa : 1	4 Sciences info Economie Sciences education Psychologie	Dans la pratique , elle a porté sur <b>une période d'observation</b> d' une dizaine de jours , entre le 7 et le 17 mars 2011 .	author => Franck Rebillard  title => Réseaux
DET + N1 pluriel	DET Périodes d'observation	5 Des : 5	1 Sciences education	Pour vérifier la cohérence des dimensions , nous avons examiné la correspondance entre les trois périodes et les indices globaux calculés pour l' ensemble <b>des périodes d'observation</b> .	author => Pierre Roque  title => Revue française de pédagogie
N2 avec déterminant	Période(s) de DET observation (s)	0	0	-	-

N2 pluriel	Période de(s) observations	0	0	-	-
N1 pluriel N2 pluriel	Périodes de(s) observations	0	0	-	-
N1 +ADJ1 de N2	Période ADJ de(s) observation (s)	1 <b>Deuxième</b> période d'observation	1  Sciences education	Au cours de notre <b>deuxième période d'observation</b> , au mois de janvier, nous constatons une forte diminution du nombre des injonctions magistrales, les injonctions destinées à créer les conditions de l'action didactique (comportement, attention) diminuant davantage que les injonctions didactiques (concernant la conduite de la leçon).	-
N1 de Adj2 N2	Période(s) de(s) ADJ observation (s)	0	0	-	-
N1 de N2 Adj2	Periode(s) de(s) observation (s) ADJ	0	0	-	-

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Point de départ	71	10  Linguistique Psychologie	Les trois articles peuvent servir de <b>point de départ</b> de cette piste de recherche.	-



			Anthropologie Histoire Sciences info Sociologie Economie Sciences Politiques Géographie Sciences education		
N1 pluriel	Points de départ	7	5  Psychologie Géographie Linguistique Sciences politiques Sciences info	A partir des sélections de direction opérées par les sujets en fonction des <b>points de départ</b> , il était possible d' examiner si les conditions d' apprentissage les avaient amenés à inverser le principe du SRP .	
DET + N1 singulier	DET Point de(s) départ(s)	43 Le : 21 Un : 17 son : 3 ce : 1 leur : 1	9  Linguistique Psychologie Histoire Anthropologie Sciences politiques Sociologie Economie Géographie Sciences info	Cependant , elle nous semble constituer <b>un point de départ</b> pertinent ( tant d' un point de vue empirique que théorique ) pour élaborer un programme de recherche dans cette voie .	
DET + N1 pluriel	DET Points de(s) départ(s)	5 Des : 5	1  Sciences education	Pour vérifier la cohérence des dimensions , nous avons examiné la correspondance entre les trois périodes et les indices globaux calculés pour l' ensemble <b>des périodes d'observation.</b>	author => Pierre Roque  title => Revue française de pédagogie
N2 avec déterminant	Point(s) de DET départ(s)	0	0		
N2 pluriel	Point de(s) départs	0	0		

N1 pluriel N2 pluriel	Points de(s) départ(s)	0	0	-	-
N1 +ADJ1 de N2	Point ADJ de(s) départ(s)	1 Un <b>bon</b> point de départ	2 Linguistique Anthropologie	En somme , l' analogie de la monnaie et du langage , si elle constitue <b>un bon point de départ</b> , doit être refondue pour précipiter des effets intéressants .	author => Jean Lassègue title => L'Homme 192
N1 de Adj2 N2	Point(s) de(s) ADJ départ(s)	0	0	-	-
N1 de N2 Adj2	Point(s) de(s) départ(s) ADJ	7 Point de départ <b>solide</b> Point de départ <b>stable</b> , Point de départ <b>difficile</b> Point de départ <b>distinct</b> Point de départ <b>absolu</b> Point de départ <b>pertinent</b>	3 Linguistique Economie Géographie	Cependant , elle nous semble constituer <b>un point de départ pertinent</b> ( tant d' un point de vue empirique que théorique ) pour élaborer un programme de recherche dans cette voie .	-

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	processus de construction	48	9 Sciences info Sciences education Sciences politiques Géographie Histoire Linguistique Psychologie Anthropologie	L' objectif de cet article n' est pas d' exposer l' intégralité du <b>processus de construction</b> sociale de cette politique publique mais bien de comprendre les modalités de pénétration de la logique gestionnaire dans le	author => François-Joseph Daniel title => Politix 91

			Sociologie	dispositif de préservation de la nature .	
N1 pluriel	Processus de construction	8	4 Sciences info Géographie Sciences education Sociologie  Psychologie Géographie Linguistique Sciences politiques Sciences info	Les <b>processus de construction</b> de l' amitié sont tout aussi sélectifs que complexes , qu' ils soient réels ou virtuels , la confiance étant une condition nécessaire à son accomplissement .	author => Virginie Lethiais  title => Réseaux 164
DET + N1 singulier	DET Processus de(s) construction(s)	36 Le : 22 Un : 12 Ce : 2	8 Sciences info Sciences politiques Sciences education Géographie Histoire Linguistique Psychologie Sociologie	Cependant , elle Notre travail centré sur <b>le processus de construction</b> d' une culture professionnelle de l' information , interrogeait notamment la question de la médiation documentaire en tentant de dégager des dominantes d' action au cours des observations et des entretiens menés sur une période de trois années et demie , de manière longitudinale .	author => Vincent Liquète  title => Les Enjeux de l'information et de la communication Dossier 2010
DET + N1 pluriel	DET Processus de(s) départ(s)	8 Les : 6 Des : 1 Ces : 1	4 Sciences info Géographie Sciences education Sociologie	Ils ont été conduits par le chercheur avec chaque enseignant pour prendre en compte <b>les processus de construction</b> de l' identité professionnelle et la manière dont s' opère l' entrée dans le métier : le premier en décembre 2004 , le second en mars 2005 , le troisième en juin 2005 .	-
N2 avec déterminant	Processus de DET construction(s)	0	0	-	-

N2 pluriel	Processus de(s) constructions	0	0	-	-
N1 pluriel N2 pluriel	Processus de(s) constructions	0	0	-	-
N1 +ADJ1 de N2	Processus ADJ de(s) construction	13 Processus <b>cyclique</b> de construction Processus <b>historique</b> de construction Processus (plus) <b>généraux</b> de construction Processus <b>dynamique</b> de construction Processus <b>interactif</b> de construction Processus <b>psychosocial</b> de construction	6 Sciences info Géographie Sciences education Sciences politiques Linguistique Psychologie	Quelques éléments de rappel permettront de donner , ici , de rapides précisions sur <b>le processus historique de construction</b> de ces mouvements .	author => Camille Goirand  title => Revue française de science politique Vol. 60
N1 de Adj2 N2	Processus de(s) ADJ construction	0	0	-	-
N1 de N2 Adj2	Processus de construction ADJ	10 Processus de construction <b>sociale</b>	5 Géographie Sciences education Sciences politiques	L' enseignant débutant est engagé dans <b>un processus de construction identitaire</b> où il revoit ses priorités	-

		Processus de construction <b>identitaire</b> Processus de construction <b>nationale</b> Processus de construction <b>territoriale</b> Processus de construction <b>relationnelle</b> Processus de construction <b>administrative</b> Processus de construction <b>discursive</b>	Histoire Sciences info	et négocie avec les différentes contraintes propres à son contexte d'insertion .	
	<b>N prép N</b>	<b>Nombre total d'occurrences</b>	<b>Dispersion</b>	<b>Exemple</b>	<b>Métadonnées</b>
Sans contraintes	Programme de recherche	74	10  Sociologie Linguistique Sciences politiques Psychologie Sciences info Géographie Economie Anthropologie Sciences education Histoire	La problématique de notre article s'inscrit dans <b>un programme de recherche</b> récent qui vise à étudier le rôle transactionnel de la monnaie à l'aide du modèle de prospection en présence d'une distribution des encaisses monétaires .	-
N1 pluriel	Programmes de recherche	6	4  Sciences info Sociologie Anthropologie Géographie	Les <b>processus de construction</b> de l'amitié sont tout aussi sélectifs que complexes , qu'ils soient réels ou virtuels , la confiance étant une condition nécessaire à son accomplissement .	author => Virginie Lethiais  title => Réseaux 164
DET + N1 singulier	DET Programme de(s) recherche(s)	53  Un : 30 Notre : 8 Le : 7	10  Sociologie Linguistique Sciences politiques	<b>Notre programme de recherche</b> trouve principalement son intérêt dans cet aspect puisqu'il s'agit de	-

		Ce : 5 Leur : 2 Son : 1	Psychologie Sciences info Géographie Economie Anthropologie Sciences education Histoire	poursuivre en France ce travail initié aux États-Unis .	
DET + N1 pluriel	DET Programmes de recherche	2  Des : 2	2  Anthropologie Sciences info	Des agences de renseignements officielles ou des officines d'espionnage préparaient en sous-main des interventions menées à l'étranger sur la base d'informations diversifiées , de fiches documentaires ou de rapports s' appuyant parfois sur <b>des programmes de recherche.</b>	author => Jackie Assayag  title => L'Homme 187-188
N2 avec déterminant	Programme de DET recherche	1  Programme (incitatif) <b>de la</b> recherche	1  Sciences education	Les données que nous utilisons ont été recueillies dans le cadre d' une recherche plus large ( Sensevy , 2007b ) , qui répond à un appel à projet , « contexte sociaux des apprentissages » , lancé en mai 2003 par <b>le programme incitatif de la recherche</b> en éducation et formation ( PIREF ) .	author => Dominique Forest  title => Revue française de pédagogie
N2 pluriel	Programme de(s) recherches	5	3  Anthropologie Sociologie Psychologie	Pour approfondir cette question importante dans l' école et pour l' école , la construction d' <b>un programme de recherches</b> en anthropologie de l' enfant ne pourrait -elle pas apporter un éclairage innovant ?	author => Marie-Armelle Barbier-Le Déroff  title => Ethnologie française Vol. 37

N1 pluriel N2 pluriel	Programmes de(s) recherches	1	1 Sociologie	Comme le montre sans détour l'existence <b>des programmes de recherches</b> financés par l' ANR qui visent à encourager la conception de dispositifs multimodaux de détection de comportements qualifiés d' « anormaux » ou de « suspects » ( algorithmes de reconnaissance comportementaux ) .	author=> Gérard Dubey  title => Cahiers internationaux de sociologie 125
N1 +ADJ1 de N2	Programme ADJ de(s) recherche(s)	6 Un <b>véritable</b> programme de recherche Un <b>tel</b> programme de recherche Le programme <b>iniciatif</b> de la recherche Un <b>important</b> programme de recherche Un programme <b>international</b> de recherche	6 Sciences education Sciences politiques Linguistique Psychologie Sociologie Economie	Ainsi , la pragmatique linguistique avait devant elle <b>un véritable programme de recherche</b> , qui d'une part permettait à la linguistique de s' affranchir de la domination formaliste en syntaxe et en sémantique et permettait de concilier deux thèses philosophiques ( anti-descriptiviste et conventionaliste ) dans un programme riche en perspective .	-
N1 de Adj2 N2	Programme (s) de(s) ADJ recherche(s)	0	0	-	-
N1 de N2 Adj2	Programme de recherche ADJ	4 Programme de recherche <b>proche</b> Programme de recherche <b>chomskien</b> Programme de recherche <b>encyclopédique</b> Programme de	4 Géographie Histoire Linguistique Sociologie	En 1784 , enfin , le même Arthaud et quelques autres donnent naissance à une véritable société savante coloniale , le Cercle des Philadelphes , qui affiche <b>un programme de recherche encyclopédique</b> , prioritairement centré sur les colonies antillaises .	author => François Regourd  title => Revue d'histoire moderne et contemporaine 55-2

		recherche <b>sociologique</b>			
--	--	----------------------------------	--	--	--

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	recherche en science	67	7 Sciences politiques Sociologie Sciences info Histoire Anthropologie Sciences education Linguistique	Le fait que ce modèle de sense - making organisationnel ne soit pas davantage mis à profit dans <b>les recherches en sciences</b> de l'information est peut-être dû au fait que les chercheurs doivent encore en évaluer la portée dans les études de comportements informationnels qui , souvent , mettent l'accent sur les composantes individuelles .	author => Dominique Maurel title => Études de communication 35
N1 pluriel	Recherches en sciences	11	6 Sciences info Sociologie Anthropologie Linguistique Sciences education Sciences info	Ces questions sont peu abordées par les sociologues et par <b>les recherches en science</b> de l'éducation : pour les premiers , les pratiques numériques des adolescents semblent propres au monde juvénile , extra- scolaires , voire a - scolaires ; les secondes se centrent sur les usages scolaires et la place des TIC dans l'éducation .	author => Cédric Fluckiger title => Revue française de pédagogie
DET + N1 singulier	DET Recherche en science	47 La : 46 Une : 1	5 Sciences politiques Sociologie Histoire Sciences info Anthropologie	La police peine en France à se voir reconnaître comme un objet commun de <b>la recherche en science</b> politique	author => Pierre Favre title => Revue française de science politique Vol. 59
DET + N1 pluriel	DET Recherches en science(s)	9 Les : 7 Nos : 1 Des : 1	5 Sciences info Sociologie Anthropologie Sciences education	La masse est toujours un paramètre important <b>des recherches en sciences</b> sociales , même lorsqu'elles ne mettent en jeu que des quantifications.	author => Michel Grossetti title => L'Année sociologique Vol. 56



			Sciences politiques		
N2 avec déterminant	Recherche en science(s) DET	0	0	-	-
N2 pluriel	Recherche en sciences	50	6 Sciences politiques Sociologie Histoire Sciences info Anthropologie Sciences education	Il faut y lire premièrement les effets de la concurrence internationale , plus fortement ressentie dans le domaine de <b>la recherche en sciences.</b>	author => Jérôme Aust title => Revue française de science politique Vol. 59
N1 pluriel N2 pluriel	Recherches en sciences	8	5 Sciences info Sociologie Anthropologie Linguistique Sciences politiques	Cette idée de la part réelle du sujet qui devrait faire l' objet <b>des recherches en sciences</b> humaines peut être dégagée chez d'autres auteurs comme par exemple le sociologue Alfred Schütz .	-
N1 +ADJ1 de N2	Recherche(s) en science(s) ADJ	4 Recherches <b>empiriques</b> en sciences Recherches <b>récentes</b> en sciences Recherche <b>interdisciplinaire</b> en science Recherche <b>fondamentale</b> en science	3 Sciences education Sciences politiques Sciences info	<b>Les recherches empiriques en sciences</b> sociales montrent que le mérite est davantage valorisé par les individus qui ont un statut social supérieur que par les individus qui ont un statut social inférieur [ 20 ] .	author => Frédéric Gonthier title => Revue française de science politique
N1 de Adj2 N2	Recherche en ADJ	0	0	-	-

	science(s)				
N1 de N2 Adj2	Recherche(s) en science(s) ADJ	73 Recherche en science sociale Recherche en science humaine Recherche en science cognitive Recherche en science politique	8 Sciences politiques Sociologie Sciences info Histoire Anthropologie Sciences education Linguistique Economie	La masse est toujours un paramètre important des <b>recherches en sciences sociales</b> , même lorsqu'elles ne mettent en jeu que des quantifications.	author => Michel Grossetti title => L'Année sociologique Vol. 56

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	Réponse à une question	15	6 Linguistique Sciences education Sciences info Géographie Histoire Sociologie	On retrouve ici l' analyse de Frank ( 1990 ) selon laquelle la <b>réponse à une question</b> rhétorique est une inférence sur une réponse négative	-
N1 pluriel	Réponses à une question	3	3 Géographie Sciences education Sciences info	Elle leur apporte des <b>réponses à des questions</b> qu' ils ne s' étaient pas même posées .	author => Ludivine Bantigny title => Revue française de pédagogie
DET + N1 singulier	DET Réponse à une question	34 La : 26 Une : 8	8 Sociologie Linguistique Psychologie Sciences info Géographie	Une partie de <b>la réponse à cette question</b> se trouve dans l' identification des officiers de l' église qui étaient chargés de montrer , ou d' organiser	author => François-Joseph Ruggiu title => Revue historique

			Economie Sciences education Histoire	la visite , d' un certain nombre de ses parties .	
DET + N1 pluriel	DET Réponses à une question	16 Les : 11 Des : 4 Ses : 1	6 Sciences education Sciences info Sociologie Economie Linguistique Psychologie	<b>Les réponses à ces questions</b> sont enregistrées à l'aide d'une échelle de type Likert allant de 1 " jamais " à 6 " presque tout le temps " .	-
N2 avec déterminant	Réponse(s) à DET une question (s)	57 Réponse à <b>cette</b> question : 24 Réponse à <b>la</b> question : 15 Réponse à <b>une</b> question : 15 Réponse à <b>notre</b> question : 2 Réponse à <b>chaque</b> question : 1	8 Sociologie Linguistique Sciences education Economie Sciences info Géographie Psychologie Histoire	Plus <b>la réponse à cette question</b> est positive , et plus les grands concepts que sont l' intégration et la socialisation sont mis en cause , tandis que leur usage devient idéologique et répressif .	author => Michel Wieviorka title => Cahiers internationaux de sociologie 125
N2 pluriel	Réponse à DET questions	3	3 Linguistique Psychologie Sciences politiques	Pour apporter des éléments de <b>réponse aux questions</b> que pose cet écart , il faut se tourner vers les dossiers disciplinaires et les ouvrir , pour à la fois suivre leur mode de fabrication et décrire les mécanismes qui se donnent à lire à travers eux , avant de nous intéresser aux instances de prise de décision des sanctions .	author => Cédric Moreau de Bellaing title => Politix 87
N1 pluriel N2 pluriel	Réponses à DET questions	23	8 Sciences education Sciences info Economie Linguistique	La manière de représenter <b>les réponses à nos trois questions</b> permet de faire un inventaire des différents types de théories pragmatiques .	-

			Sociologie Géographie Psychologie Sciences politiques		
N1 +ADJ1 de N2	Réponse(s) ADJ à DET question(s)	16 Réponse <b>positive...</b> Réponse <b>assurée...</b> Réponse <b>possible</b> Réponse <b>claire</b> Réponse <b>universelle</b> Réponse <b>dissonante</b> Réponse <b>simple</b> Réponse <b>univoque</b> Réponse <b>cohérente</b>	6 Sciences education Sciences politiques Linguistique Psychologie Sociologie Economie	La naissance du mouvement de défense des droits civiques donne <b>une réponse claire à cette question .</b>	author => Lun Zhang title => Cahiers internationaux de sociologie 122
N1 de Adj2 N2	réponse(s) à DET ADJ question(s)	0	0	-	-
N1 de N2 Adj2	Réponse(s) à DET question(s) ADJ	0	0	.	-

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	résultat de l'étude	43	8 Sciences education Psychologie Economie Sciences politiques Sciences info Sociologie Géographie Anthropologie	De fait , <b>les résultats de l' étude</b> présentée ici militent plutôt en faveur de l' absence de schéma de la réussite professionnelle pour les demandeurs d' emploi .	author => A. Somat title => Le travail humain

N1 pluriel	résultats de l'étude	40	7 Sciences education Psychologie Economie Sciences politiques Sciences info Sociologie Géographie	La place limitée ne nous a pas permis de développer <b>les résultats de l'étude</b> ( ni de spécifier l' impact des variables en termes d' intensité d' effet ) mais cette limitation est en un sens éclairante .	-
DET + N1 singulier	DET résultat de l'étude	1 Le : 1	1 Psychologie	<b>Le résultat</b> principal de notre <b>étude</b> est de montrer que , bien que les conducteurs âgés aient un risque majeur d' être impliqué dans un accident mortel dans ces situations de traversée de voie(s) , ils sont néanmoins moins dangereux pour les autres impliqués par rapport à des conducteurs plus jeunes , lorsqu' on évalue la dangerosité par le nombre d' avp .	author => S. Lafont title => Le travail humain
DET + N1 pluriel	DET résultats de l'étude	31 Les : 28 Ces : 1 Quelques : 1 Des : 1	6 Psychologie Sciences education Géographie Sciences politiques Sociologie Histoire	D'après <b>les résultats des études</b> menées avec le paradigme d' amorçage , la préactivation d' un stimulus qui partage la première syllabe avec le mot cible affecte le traitement de ce mot cible ( Carreiras & Perea , 2002 ; Dominguez , de Vega & Cuetos , 1997 ) .	-
N2 avec déterminant	Résultat(s) de DET étude(s)	41 Résultat de cette étude : 15 Résultat de l'étude : 9 Résultat des études : 8 Résultat d'une étude : 6 Résultat de notre étude : 2 Résultat de plusieurs études : 1	8 Psychologie Sciences education Economie Sociologie Géographie Sciences politiques Histoire Sciences info	Cet article souhaite contribuer à l' analyse de ces particularités en présentant <b>quelques résultats d' une étude</b> de terrain menée sur les structures territoriales contemporaines de ce plateau des Boloven [ 2 ] en comparaison avec celle des plateaux du Centre Viêt-nam à propos de la caféiculture ( Fortunel , 2003 ) .	author => Frédéric Fortunel title => L'Espace géographique
N2 pluriel	résultat de DET études	8	3	<b>Les résultats des études</b> menées dans divers pays et	author => Laëtitia

		Des : 8	Sciences education Economie Psychologie	à différents niveaux scolaires révèlent que cette acquisition est loin d'être triviale .	Pinet title => Revue française de pédagogie
N1 pluriel N2 pluriel	Résultats de(s) études	12	4 Sciences education Economie Psychologie Sciences politiques	Dans l' année , on débat <b>des résultats d' études</b> qui sont présentés , c' est très intéressant , il faut le reconnaître , pour soi. Mais après. le problème c' est qu' après il faut se dépêcher de faire le rapport .	author => Yves Lochard title => Politix
N1 +ADJ1 de N2	Résultat(s) ADJ de(s) étude(s)	6	3 Psychologie Economie Histoire	Face au manque de <b>résultats significatifs des premières études</b> empiriques , un nombre croissant d' auteurs ont mis l' accent sur une série de problèmes méthodologiques susceptibles d' invalider les relations testées .	
N1 de Adj2 N2	Résultat(s) de(s) ADJ étude(s)	4	2 Economie Sciences education	Face au manque de <b>résultats significatifs des premières études</b> empiriques , un nombre croissant d' auteurs ont mis l' accent sur une série de problèmes méthodologiques susceptibles d' invalider les relations testées .	
N1 de N2 Adj2	Résultat(s) de DET étude(s) ADJ	12	3 Sciences education Economie	<b>Les résultats des études dites quasi expérimentales</b> comparant élèves présentant des difficultés	

		<b>international</b> es Résultat des études <b>empiriques</b> Résultat des études <b>antérieures</b> Résultat des études dites quasi <b>expérimental</b> es	Psychologie	dont certains doublent leur année tandis que d'autres sont promus	
--	--	---	-------------	---	--

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	recueil de données	35	7 Psychologie Sciences education Anthropologie Sciences info Géographie Sociologie Sciences politiques	La plus grande partie du travail de <b>recueil de données</b> a été consacrée à la collecte et à la numérisation de la sectorisation des collèges publics parisiens au cours de la période 1997 - 2004 .	author => Gabrielle Fack  title => Actes de la recherche en sciences sociales 180
N1 pluriel	Recueils de(s) donnée(s)	8	4 Sciences education Psychologie	Ce travail, fondé sur une revue des travaux sur le sujet et <b>les recueils de données</b> statistiques de la CNNIC [ 2 ] , s'appuie sur trois études empiriques ( Pons et al. , 2006 ; Puel , Xu , 2007 ; Puel , Zhang , 2007 ) [ 3 ] .	author => Gilles Puel  title => L'Espace géographique Vol. 38
DET + N1 singulier	DET Recueil de(s) donnée(s)	14 Le : 12 Aucun : 1 Un : 1	6 Psychologie Anthropologie Sciences education Sciences info Economie Sociologie	Le matériel utilisé pour le séquençement des présentations et <b>le recueil des données</b> est le même que celui de la première expérience .	-
DET + N1 pluriel	DET Recueils de (s) donnée(s)	4 Les : 2 Des : 2	4 Anthropologie Sciences education Géographie Psychologie	Pour ce faire, nous avons pris appui sur <b>les</b> deux premiers <b>recueils de données</b> : le premier rend compte des préconceptions de la relation pédagogique et le second , des contraintes issues du terrain .	-



N2 avec déterminant	Recueil(s) de DET donnée(s)	12 Recueil de ces données : 1 Recueil des données : 11	5 Psychologie Anthropologie Géographie Sciences education Sciences info	C'est pourquoi il semble important de permettre <b>le recueil de ces données</b> pour préciser davantage comment affects et cognition sont intriqués en situation de travail .	author => B. Cahour  title => Le travail humain
N2 pluriel	Recueil de(s) données	27	7 Anthropologie Sociologie Psychologie Sciences info Sciences education Géographie Sciences politiques	Très vite , les chercheurs se sont rendu compte qu' il n' était pas possible de conclure avec pareil dispositif de <b>recueil de données</b> .	-
N1 pluriel N2 pluriel	Recueils de(s) données	8	4 Sciences education Psychologie Anthropologie géographie	Mais rares sont les études qui prennent en compte l' ensemble de ces paramètres , car les difficultés à surmonter sont multiples : il s' agit dans un premier temps de synchroniser des <b>recueils de données</b> quantitatives de performance en relation avec des mesures physiologiques et des indicateurs comportementaux ( expressions faciales , directions du regard ) .	author => M. Wolff  title => Le travail humain
N1 +ADJ1 de N2	Recueil(s) ADJ de(s) donnée(s)	3 Un <b>simple</b> recueil Le <b>premier</b>	2 Sciences education Anthropologie	Contribution à l' étude archéologique d' un district saharien ( tmie , 19 , 1932 ) , quoique l' ouvrage n' ait rien d'	author => Vincent Debaene  title => L'Homme

		recueil : 2		ethnologique et se donne comme <b>un simple recueil de données</b> ( gravures rupestres , outillages lithiques , etc. ) , destiné à verser « un certain nombre de documents [ . ] au dossier des études sahariennes » ( Monod 1932 : 2 ) .	
N1 de Adj2 N2	Recueil(s) de(s) ADJ donnée(s)	0	0	-	-
N1 de N2 Adj2	Recueil de donnée ADJ	6 Recueil de donnée <b>quantitative</b> Recueil de donnée <b>qualitative</b> Recueil de donnée <b>statistique</b> Recueil de donnée <b>visuelle</b>	4 Géographie Histoire Linguistique Sociologie	Le contexte de la prise de vue les choix du cadrage photographique et de l'angle de vue , la méthode de <b>recueil des données visuelles</b> , le rôle de la prise de vue dans la relation vécue entre l'observateur et son terrain sont rarement interrogés , commentés .	author => Sylvaine Conord title => Ethnologie française Vol. 37

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	principe d'organisation	18	6 Linguistique Sciences politiques Géographie Histoire Psychologie Science info	En dépit de la diversité des significations dont est investi <b>le principe d'organisation</b> , les vues développées dans le cadre de cette anatomie comparée - que l'on prend coutume de désigner également du nom de « sciences de l'organisation » - entraînent une série de changements convergents radicaux dans le sens assigné aux concepts taxinomiques et à la classification .	author => Dominique Guillo title => Politix

N1 pluriel	Principes de(s) organisation	11	6 Linguistique Géographie Histoire Sciences politiques Psychologie Sciences info	Ce sont ces <b>principes d'organisation</b> des événements qu' il appelle « cadre ou frame » .	author => Bernard Kalaora  title => L'Espace géographique Tome 39
DET + N1 singulier	DET Principe de(s) organisation	4 Le : 2 Ce : 1 Un : 1	2 Sciences politiques géographie	<b>Ce principe d'organisation</b> se trouve au cSur du dispositif spatial du Conservatoire du littoral lors de sa fondation : faire en sorte que la mer soit visible en aménageant des sites naturels et paysagers et en rendant possibles les accès	author => Bernard Kalaora  title => L'Espace géographique Tome 39
DET + N1 pluriel	DET Principes de (s) organisation(s)	5 Les : 4 Ces : 1	3 Histoire Sciences politiques Géographie	Dans un tel cadre , <b>les principes d'organisation</b> , de corrélation des formes ou de plan de composition dessinent un véritable langage des corps .	author => Dominique Guillo  title => Politix
N2 avec déterminant	Principe(s) de DET organisation(s)	3 Principe d' <b>une</b> organisation Principe de l'organisation	3 Linguistique Psychologie Histoire	La direction du PCF se trouve quant à elle partagée sur cette question , mais se montre cependant plutôt favorable au <b>principe d'une organisation</b> d'autodéfense spécialisée .	author => Georges Vidal  title => Revue historique 640
N2 pluriel	Principe de(s) organisations	0	0	-	-

N1 pluriel N2 pluriel	Principe(s) ) de(s) organisati on(s)	0	0	-	-
N1 +ADJ1 de N2	Principe(s) ) ADJ de(s) organisati on(s)	3  Principe <b>taxinomique</b> d'organisation Principe <b>dominant</b> de l'organisation Principe <b>favorable</b> d'une organisation	3  Histoire Psychologie Sciences politiques	La gestion et le contrôle de l' espace et du temps carcéral constituent des <b>principes dominants de</b> <b>l' organisation</b> pénitentiaire , en raison des risques sécuritaires inhérents à leur absence de couverture .	author => V. Moulin  title => Le travail humain
N1 de Adj2 N2	Principe(s) ) de(s) ADJ organisati on(s)	0	0	-	-
N1 de N2 Adj2	Principe d'organis ation ADJ	2  Principe d'organisation <b>indexicale</b> Principe d'organisation <b>pénitentiaire</b>	2  Linguistique Psychologie	La gestion et le contrôle de l' espace et du temps carcéral <b>constituent des</b> <b>principes dominants de</b> <b>l' organisation</b> <b>pénitentiaire</b> , en raison des risques sécuritaires inhérents à leur absence de couverture .	author => V. Moulin  title => Le travail humain

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	point d'appui	28	8  Linguistique Sciences politiques Géographie Histoire Psychologie	Mais elle peut être outillée car elle pourrait bien être ambiguë et ne déjouer localement les disciplines qu' en renforçant leur <b>point d'</b> <b>appui</b> stratégique , l' emprisonnement du	-

			Sociologie Anthropologie Sciences education	corps dans l' âme , principe de toutes les docilités , et des plus puissantes parce que les moins apparentes	
N1 pluriel	Points de(s) appui	17	6  Géographie Histoire Psychologie Sciences politiques Sciences education Sociologie	Nous le constaterons et nous montrerons comment , dans ce contexte , les villes sahéliennes peuvent être des <b>points d' appui</b> au développement dans des environnements marqués par la pauvreté et l' instabilité .	author => Anne Ouallet  title => Annales de géographie
DET + N1 singulier	DET Point de(s) appui	9  Un : 6 Aucun : 1 Le : 1 Leur : 1	7  Histoire Sciences politiques Anthropologie Géographie Linguistique Sciences education Sociologie	La mission de Lagrené est jugée de première importance : il doit trouver <b>un point d' appui</b> , obtenir les mêmes avantages que les Britanniques , découvrir de quoi inciter les armateurs français à se lancer dans l' aventure chinoise , sans parler de la question religieuse , dont la Monarchie de Juillet souhaite relever la bannière .	author => Guy DURAND  title => Revue d'histoire moderne et contemporaine 57-1
DET + N1 pluriel	DET Points de (s) appui (s)	6  Les : 2 Des : 2 Ces : 1 Quelques : 1	4  Géographie Histoire Sciences education Sciences politiques	D'où l' intérêt de trouver <b>les points d' appuis</b> développés par Guizot .	author => Guy DURAND  title => Revue d'histoire moderne et contemporaine 57-1
N2 avec déterminant	Point(s) de DET appui(s)	0	0	-	-

N2 pluriel	Point de(s) appuis	0	0	-	-
N1 pluriel N2 pluriel	Points de(s) appuis	3	1 Histoire	Cet ensemble , complété par quelques modestes <b>points d' appuis</b> sur les côtes de l' Afrique de l' ouest , de l' Océan Indien ou de la Guyane , est alors progressivement rattaché à l' autorité directe de la monarchie .	author => François Regourd  title => Revue d'histoire moderne et contemporaine 55-2
N1 +ADJ1 de N2	Point(s) ADJ de(s) appui(s)	10 <b>Nouveau point</b> <b>Divers points</b> <b>Nombreux points</b> <b>Modestes points</b>	5  Sciences politiques Géographie Psychologie Histoire Sociologie	Ces capitales régionales , anciennes capitales d' Empire sont <b>les nouveaux points d' appui</b> du développement dans le nord du Mali .	author => Anne Ouallet  title => Annales de géographie
N1 de Adj2 N2	Point(s) de(s) ADJ appui(s)	0	0	-	-
N1 de N2 Adj2	Point d'appui ADJ	5  Point d'appui <b>stratégique</b>	4  Sociologie Géographie	Et le nouveau « mille - feuille » de la gouvernance portuaire toulonnaise lui offrira	author => Jean Debrie  title => Annales de

	<b>N prép N</b>	<b>Nombre total d'occurrences</b>	<b>Dispersion</b>	<b>Exemple</b>	<b>Métadonnées</b>
Sans contraintes	partie de l'article	23	7 Sciences info Economie	Ceci est le matériel analysé dans la première <b>partie de l'article.</b>	author => Jairo Ferreira title => Les Enjeux

		Point d'appui <b>privilegié</b> Point d'appui <b>heuristic</b> Point d'appui <b>insitutionnel</b> Point d'appui <b>juridique</b>	Sciences education Sciences politiques	<b>un point d'appui</b> <i>heuristique</i> particulièrement éclairant .	géographie 669
--	--	---	---	---	----------------

			Linguistique Sciences politiques Géographie Psychologie Sciences education		de l'information et de la communication Volume 2010
N1 pluriel	Parties de l' article	0	0	-	-
DET + N1 singulier	DET partie de l'article	23  La : 20 Une : 3	7  Sciences info Géographie Economie Sciences education Linguistique Psychologie Sciences politiques	<b>La première partie de</b> cet <b>article</b> posera plus précisément les questions qui nous intéressent ici et le cadre théorique dans lequel elles s'intègrent .	author => Virginie Lethiais  title => Réseaux 164
DET + N1 pluriel	DET parties de l'article	0	0	-	-



N2 avec déterminant	Partie de article DET	23 Partie de <b>cet</b> article : 13 Partie de l'article : 9 Partie <b>des</b> articles : 1	7 Sciences info Géographie Economie Sciences education Linguistique Psychologie Sciences politiques	Une <b>partie des articles</b> de notre échantillon se situe clairement dans le compte rendu factuel de la compétition électorale .	author => Annelise Touboul title => Réseaux
N2 pluriel	Partie de(s) articles	1	1 Sciences info	Une <b>partie des articles</b> de notre échantillon se situe clairement dans le compte rendu factuel de la compétition électorale .	author => Annelise Touboul title => Réseaux
N1 pluriel N2 pluriel	Parties de(s) articles	0	0	-	-
N1 +ADJ1 de N2	Partie(s) ADJ de(s) article(s)	20 La <b>première</b> partie La <b>dernière</b> partie La partie <b>précédente</b> La partie <b>empirique</b>	7 Sciences info Sciences politiques Géographie Psychologie Histoire Sociologie Linguistique	En effet , globalement la connaissance est un facteur déterminant dans le choix exprimé par les experts comme nous le montrerons dans <b>la partie empirique de l' article .</b>	-
N1 de Adj2 N2	Partie(s) de(s) ADJ article(s)	0	0	-	-

N1 de N2 Adj2	Partie de l'article ADJ	0	0	-	-
------------------	----------------------------	---	---	---	---

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contrain tes	outil de communication	30	8  Sciences info Economie Linguistique Sciences politiques Géographie Psychologie Sciences education Sociologie	En deuxième lieu , les personas constituent également <b>un outil de communication</b> et de collaboration efficace , en permettant par exemple de véhiculer des scénarios d' utilisation associés à une certaine représentation d'utilisateur .	author => É. Brangier  title => Le travail humain
N1 pluriel	Outils de communication	21	5  Sciences info Sciences politiques Sciences education Histoire Sociologie	On peut aussi signaler le développement des <b>outils de communication</b> propres des petits groupes militants , associatifs ou culturels .	author => Cégolène Frisque  title => Revue française de science politique Vol. 60
DET + N1 singulier	DET outil de communication	6  Un : 5 Le : 1	5  Sciences info Sciences education Linguistique Psychologie Sciences politiques	Ce dispositif créé en 1996 permet à la fédération de suivre de près l' évolution des aspirations des familles et constitue <b>un outil de communication</b> interne pour les dirigeants de la PEEP .	author => Philippe Gombert  title => Revue française de pédagogie

DET + N1 pluriel	DET outils de communication	12  Les : 6 Des : 3 Ces : 1 Leurs : 1	4  Sciences info Sciences education Histoire Sciences politiques	<b>Les outils de communication</b> comme la messagerie instantanée ou les blogs permettent aux adolescents de poursuivre à domicile les relations avec leurs amis , camarades de classe ou copains du quartier , en échappant , ne serait -ce que partiellement , à la surveillance familiale ( Pasquier , 2005 ) .	author => Cédric Fluckiger  title => Revue de française de pédagogie
N2 avec déterminant	Outil de DET communication	0	0	-	-
N2 pluriel	Outil de(s) communications	0	0	-	-
N1 pluriel N2 pluriel	Outils de(s) communications	0	0	-	-
N1 +ADJ1	Outil(s) ADJ de(s) communication(s)	4	3	La mise en place d' <b>un nouvel outil de</b>	author => Émilie

de N2		<b>Meilleur</b> outil <b>Nouveaux</b> outils <b>Simple</b> outil	Sciences politiques Sciences info Linguistique	<b>communication</b> , à savoir le microblogging , peut considérablement améliorer la circulation de l' information .	Ogez title => Documentaliste- Sciences de l'Information Vol. 47
N1 de Adj2 N2	Outil(s) de(s) ADJ communication(s)	0	0	-	-
N1 de N2 Adj2	Outil de communication ADJ	5 Outil de communication n <b>externe</b> Outil de communication n <b>interne</b> Outil de communication n <b>numérique</b> Outil de communication n <b>synchrone</b> Outil de communication n <b>susceptible</b>	3 Sciences politiques Sciences info Sciences education	<b>Comme outil de communication susceptible</b> de développer , nous allons le voir , des communautés , « such systems cannot be developed in small ways on small machines .	author => Pascal Robert title => Etudes de communication

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	objet d'analyse	39	10 Sciences info Economie Linguistique Sciences politiques Géographie Psychologie Sciences education Sociologie Anthropologie Histoire	En prenant pour <b>objet d'analyse</b> les programmes de cours de français au premier degré secondaire , on montre ensuite que les différents réseaux d'enseignement ne se sont pas lancés dans ce mouvement pédagogique avec la même intensité .	-
N1 pluriel	Objets d'analyse	1	1 Anthropologie	L'ensemble des fiches consacrées partiellement ou totalement aux instruments de musique renvoie donc l' image d' une « organographie » assez complète , associée à un ensemble plus vaste et reliée à d'autres <b>objets d' analyse</b> ( les circonstances du jeu renvoient par exemple aux rites ) , mais qui laisse par ailleurs entrevoir une organologie assez systématique ( classification , répartition géographique , etc. ) .	author => Brice Gérard  title => L'Homme 191
DET + N1 singulier	DET d'analyse Objet	31 Le (l') : 29 Un : 2	10 Sciences info Economie Linguistique Sciences politiques Géographie Psychologie Sciences education Sociologie Anthropologie Histoire	Ce corpus a fait <b>l'objet d'une analyse</b> de contenu thématique .	author => Marie-France Peyrelong  title => Communication & Organisation 36
DET + N1 pluriel	DET d'analyse objets	1 <b>D'autres</b>	1 Anthropologie	L'ensemble des fiches consacrées partiellement ou totalement aux instruments de musique renvoie donc l' image d' une « organographie » assez complète , associée à un ensemble plus vaste et reliée à <b>d'autres objets</b>	author => Brice Gérard  title => L'Homme 191

				<b>d' analyse</b> ( les circonstances du jeu renvoient par exemple aux rites ) , mais qui laisse par ailleurs entrevoir une organologie assez systématique ( classification , répartition géographique , etc. ) .		
N2 avec déterminant	Objet de analyse	DET	20  Objet d'une analyse : 14 Objet de notre analyse : 3 Objet de l'analyse : 1 Objet de mon analyse : 1 Objet de son analyse : 1	8  Psychologie Sciences education Sciences politiques Sociologie Linguistique Anthropologie Histoire Sciences info	Les données font <b>l'objet d' une analyse</b> quantitative ( structure des descriptions ) et qualitative ( symbolique de l' espace vécu ) .	-
N2 pluriel	Objet analyses	de(s)	9	6  Psychologie Sciences education Sciences politiques Sociologie Linguistique Anthropologie	Enfin , les matrices de corrélations ont fait <b>l'objet d' analyses</b> en composantes principales et en facteurs communs afin de déterminer si la part de variance expliquée par la première composante ou par le premier facteur est supérieure pour les participants crédités des notes les plus faibles .	-
N1 pluriel N2 pluriel	Objets analyses	de(s)	0	0	-	-

N1 +ADJ1 de N2	Objet(s) ADJ de(s) analyse(s)	1  Objet <b>central</b> de	1  Anthropologie	Ils composent <b>l'objet central de mon analyse</b> .	author => Jean-Sébastien Eideliman  title => Ethnologie française Vol. 39
N1 de Adj2 N2	Objet(s) de(s) ADJ analyse(s)	3  Objet de <b>nombreuses</b> analyses	2  Economie Sciences info	La reconnaissance , d'abord pensée comme concept philosophique , a fait <b>l'objet de nombreuses analyses</b> psychosociologiques , en particulier pour souligner les attentes des personnels au travail .	author => Valérie Lépine  title => Communication & Organisation 36
N1 de N2 Adj2	Objet d'analyse ADJ	22  Objet d'analyse (assez) <b>différent</b> Objet d'une analyse <b>détaillée</b> Objet d'une analyse <b>automatique</b> Objet d'une analyse <b>quantitative</b> Objet d'une analyse <b>factorielle</b> Objet d'une analyse <b>narrative</b> Objet d'une analyse <b>qualitative</b> Objet de (nombreuses) analyses	9  Sciences info Economie Linguistique Sciences politiques Géographie Psychologie Sciences education Sociologie Histoire	Pour chaque concepteur et pour les deux premières phases ( sans et avec personas ) , les données ont fait <b>l'objet d' une analyse quantitative et qualitative</b> comparative	author => É. Brangier  title => Le travail humain

		<b>psychosociologiques</b> Objet d'analyse <b>géographique</b> Objet d'analyse <b>sociologique</b> Objet d'analyse <b>administrative</b> Objet d'une analyse <b>modeste</b> Objet d'analyse <b>linguistique</b> Objet d'analyse <b>globale</b>			
--	--	--	--	--	--



	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	mode d'action	39	7 Sciences politiques Psychologie Sciences education Sociologie Anthropologie Histoire Sciences education	Le <b>mode d' action</b> patronal traditionnel de la bienfaisance ne va désormais plus de soi .	author => Philippe Hamman  title => Politix 84
N1 pluriel	Modes d'action	32	6	<b>Les modes d'action</b> de l'école en matière culturelle sont variés , des plus directs aux plus indirects .	author => Sylvie Octobre  title => Revue française de pédagogie
DET + N1 singulier	DET mode d'action	5 Le : 3 Ce : 1 Son : 1	3 Sciences politiques Sciences education Histoire	Il est donc possible de parler d' un habitus du héraut , dès lors qu' on relève une telle adéquation entre son apparence corporelle et vestimentaire , <b>son mode d' action</b> spécifique et l' évaluation sociale des qualités que ce mode d' action met en jeu .	author => Catherine Goblot-Cahen  title => Revue historique
DET + N1 pluriel	DET modes d'action	20 Les : 7 Des : 5 Ces : 4 D'autres : 2 Ses : 2	6 Sciences politiques Sciences education Histoire Psychologie Sociologie Anthropologie	Les phénomènes sociaux humains n' émergent pas d' une interaction entre des individus dont les buts et <b>les modes d' action</b> seraient préprogrammés .	author => Jean Lassègue  title => L'Homme 192

N2 avec déterminant	Mode de DET action	0	0	-	-
N2 pluriel	Mode de(s) actions	1	1 Sciences politiques	Comme le suggèrent ces pistes , on gagne à penser ces propositions de réformes sur <b>le mode d' actions</b> visant à instituer de nouvelles lignes de partage dont l'objectif pourrait bien être la défense et la restauration de l'illusion de la pureté au sein de la sphère politique .	author => Éric Phélippeau  title => Revue française de science politique Vol. 60
N1 pluriel N2 pluriel	Modes de(s) actions	1	1 Sciences politiques	La période 1997 - 2002 constitue le moment où un ensemble de mesures touchant , d'une part , à l'organisation interne , d'autre part , aux <b>modes d' actions</b> de la police et , enfin , aux relations entre la police et la population sont formulées au plan national sous la dénomination de « police de proximité » .	author => Sebastian Roché  title => Revue française de science politique Vol. 59
N1 +ADJ1 de N2	Mode(s) ADJ de(s) action(s)	2 <b>Nouveau</b>	1 Sciences politiques	C'est ensuite moins en « inventant » de <b>nouveaux modes d'</b>	author => Jérôme Aust

		mode d'action		<b>action</b> qu' en étendant et en radicalisant des registres d' action historiquement utilisés que le ministère de l' Enseignement supérieur et de la Recherche la conduit .	title => Revue française de science politique Vol. 59
N1 de Adj2 N2	Mode (s) de(s) ADJ action(s)	0	0	-	-
N1 de N2 Adj2	Mode d'action ADJ	2  Mode d'action <b>spécifique</b> Mode d'action <b>politique</b>	2  Sciences politiques Histoire	De façon générale , derrière la cohésion apparente d' un mouvement social , les motivations , les représentations et les comportements collectifs sont marqués par le caractère hétérogène , voire contradictoire <b>des modes d' action politique</b> , en particulier au sein des milieux populaires [ 35 ] .	author => Camille Goirand  title => Revue française de science politique Vol. 60

	N prép N	Nombre total d'occurrences	Dispersion	Exemple	Métadonnées
Sans contraintes	ensemble de pratique	34	7  Sciences politiques Sciences education Sciences info Sociologie Anthropologie Histoire Sciences education	Le contrôle et la répression sont ainsi plus grands qu' avant le début des programmes et influencent <b>l'ensemble des pratiques</b> territoriales des habitants , qu' il s' agisse d' agriculture ou de prélèvement des ressources naturelles ( et notamment la chasse coutumière ) .	author => Estienne Rodary  title => L'Espace géographique

N1 pluriel	ensembles de pratique	1	1 Anthropologie	André Schaeffner a participé aux missions Dakar - Djibouti ( 1931 - 1933 ) et Sahara - Soudan ( 1935 ) et a pu , dans ce cadre , observer les aspects variés de plusieurs <b>ensembles de pratiques</b> musicales d' Afrique .	author => Brice Gérard  title => L'Homme 191
DET + N1 singulier	DET ensemble de pratique	35 Un : 18 Le : 16 Cet : 1	7 Sciences politiques Sciences education Sciences info Histoire Anthropologie Géographie Sociologie	Dans sa première acception , la médiation est donc un <b>ensemble de pratiques</b> , qui se développent dans des domaines institutionnels différents .	author => Simon Gadras  title => Les Enjeux de l'information et de la communication Dossier 2010
DET + N1 pluriel	DET ensembles de pratique	1	1 Anthropologie	André Schaeffner a participé aux missions Dakar - Djibouti ( 1931 - 1933 ) et Sahara - Soudan ( 1935 ) et a pu , dans ce cadre , observer les aspects variés de <b>plusieurs ensembles de pratiques</b> musicales d' Afrique .	author => Brice Gérard  title => L'Homme 191

N2 avec déterminant	Ensemble de DET pratique	0	0	-	-
N2 pluriel	Mode de(s) actions	1	1 Sciences politiques	Comme le suggèrent ces pistes , on gagne à penser ces propositions de réformes sur <b>le mode d' actions</b> visant à instituer de nouvelles lignes de partage dont l' objectif pourrait bien être la défense et la restauration de l' illusion de la pureté au sein de la sphère politique .	author => Éric Phélippeau  title => Revue française de science politique Vol. 60
N1 pluriel N2 pluriel	Modes de(s) actions	1	1 Sciences politiques	La période 1997 - 2002 constitue le moment où un ensemble de mesures touchant , d'une part , à l' organisation interne , d'autre part , aux <b>modes d' actions</b> de la police et , enfin , aux relations entre la police et la population sont formulées au plan national sous la dénomination de « police de proximité » .	author => Sebastian Roché  title => Revue française de science politique Vol. 59
N1 +ADJ1 de N2	Mode(s) ADJ de(s) action(s)	2 <b>Nouveau</b>	1 Sciences politiques	C' est ensuite moins en « inventant » de <b>nouveaux modes d' action</b> qu' en	author => Jérôme Aust

		mode d'action		étendant et en radicalisant des registres d' action historiquement utilisés que le ministère de l' Enseignement supérieur et de la Recherche la conduit .	title => Revue française de science politique Vol. 59
N1 de Adj2 N2	Mode (s) de(s) ADJ action(s)	0	0	-	-
N1 de N2 Adj2	Mode d'action ADJ	2  Mode d'action <b>spécifique</b> Mode d'action <b>politique</b>	2  Sciences politiques Histoire	De façon générale , derrière la cohésion apparente d' un mouvement social , les motivations , les représentations et les comportements collectifs sont marqués par le caractère hétérogène , voire contradictoire <b>des modes d' action politique</b> , en particulier au sein des milieux populaires [ 35 ] .	author => Camille Goirand  title => Revue française de science politique Vol. 60

### Annexe 3

#### Tableau récapitulatif

Exemple	N1	déterminant	N2
<b>Critère de selection</b>	- peu de constraints	- pas de déterminant pour le N2	- est en général singulier  -un nom d'idéalité ou un nom d'activité  -un nom abstrait
<b>Analyse des données</b>	- peu de constraints	- article défini contracté	- est toujours au pluriel  -un nom abstrait
<b>Augmentation du nombre</b>	- toujours au singulier	- la forme contractée (de + le = du)	- est toujours au singulier
<b>Enquête de terrain</b>	-peu de constraints	-pas de déterminant pour le N2	- est singulier -le complément joue un rôle d'adjectif
<b>Champ d'étude</b>	-appartient à la classe sémantique de noms d'espace	-possibilité d'insérer un DET devant le N2	- est en général singulier
<b>Objectif de recherche</b>	-peu de constraints	-possibilité d'insérer un DET devant le N2	- est singulier
<b>Groupe de travail</b>	-peu de contraintes	-pas de déterminant pour le N2	- toujours singulier
<b>Objet d'étude</b>	-peu de contraintes		-peu de contraintes
<b>Période d'observation</b>	-peu de contraintes	-pas de déterminant pour le N2	-N2 toujours singulier
<b>Point de départ</b>	-peu de contraintes	-pas de déterminant pour le N2	-N2 toujours singulier
<b>Processus de construction</b>	-la même forme : pour le N1 au singulier ou ou pluriel	-pas de déterminant pour le N2	-N2 toujours singulier
<b>Programme de</b>	-peu de		-N2 toujours

<b>recherche</b>	contraintes		singulier
<b>Recherche en science</b>	-peu de contraintes	-pas de déterminant pour le N2	-peu de contraintes
<b>Réponse à une question</b>	-peu de contraintes		-peu de contraintes
<b>Résultat de l'étude</b>	-peu de contraintes		-peu de contraintes
<b>Recueil de données</b>	-peu de contraintes		-N2 toujours pluriel
<b>Principe d'organisation</b>	-peu de contraintes		-N2 toujours singulier
<b>Point d'appui</b>	-peu de contraintes	-pas de déterminant pour le N2	-peu de contraintes
<b>Partie de l'article</b>	-N1 toujours au singulier		-peu de contraintes
<b>Outil de communication</b>	-peu de contraintes	-pas de déterminant pour le N2	-N2 toujours singulier
<b>Objet d'analyse</b>	-peu de contraintes		-peu de contraintes
<b>Mode d'action</b>	-peu de contraintes	-pas de déterminant pour le N2	-peu de contraintes
<b>Ensemble de pratique</b>	-peu de contraintes	-pas de déterminant pour le N2	-peu de contraintes
<b>Diffusion d'information</b>	-N1 toujours au singulier		-peu de contraintes
<b>Valeur de référence</b>	-peu de contraintes	-pas de déterminant pour le N2	-N2 toujours singulier



## Table des matières

Remerciements .....	3
Sommaire .....	5
Introduction .....	6
<b>PARTIE 1 - NOTION DE LA COLLOCATION SOUS DIFFERENTES PERSPECTIVES .....</b>	<b>9</b>
CHAPITRE 1. INTERET DU PHENOMENE DES COLLOCATIONS .....	10
1. Etudes théoriques sur les collocations : de Bally à nos jours .....	12
1.1. L'école britannique : John Rupert Firth, John Sinclair.....	19
1.2. Propriétés syntaxiques des collocations .....	23
1.3. Structures des collocations .....	25
Types de variations morphologiques et syntaxiques .....	26
1.4. Les collocations dans les dictionnaires.....	28
Les fonctions lexicales d'Igor Mel'čuk.....	40
Les collocations dans les logiciels .....	42
Le dictionnaire des collocations en polonais.....	48
1.5. Synthèse.....	51
<b>PARTIE 2 - METHODOLOGIE ET PROBLEMATIQUE .....</b>	<b>53</b>
CHAPITRE 2. PROBLEMATIQUE : PROPRIETES SYNTAXIQUES DE COLLOCATIONS.....	54
2. Corpus et outil d'extraction .....	55
3. Le corpus Scientext des sciences humains.....	56
4. L'outil de traitement de corpus utilisé : Lexicoscope .....	57
5. La grille d'analyse proposée .....	59
<b>PARTIE 3 - ANALYSE ET PROPOSITION DE LA MODELISATION .....</b>	<b>69</b>
CHAPITRE 3. ANALYSE DES PROPRIETES SYNTAXIQUES DES COLLOCATIONS ET PISTES DE CODAGE .....	70
CHAPITRE 4. <i>PROPOSITION DE LA MODELISATION</i> .....	80
Conclusion.....	88
Bibliographie.....	91
Sigles et abréviations utilisés.....	95
Table des illustrations.....	96
Table des matières .....	158

**MOTS-CLÉS** : collocation, modélisation des collocations, propriétés syntaxiques, linguistique du corpus

### **RÉSUMÉ**

La modélisation des propriétés syntaxiques des collocations n'est pas une chose facile. Le présent travail se veut concentrer sur l'observation des constructions syntaxiques dans l'intention de comprendre leur fonctionnement, ainsi que proposer le codage de ces éléments. L'objectif final est de fournir les informations pour pouvoir créer un jour un dictionnaire électronique utilisable en ligne. Pour explorer cette problématique, nous avons recouru aux méthodes de la linguistique de corpus. Le champ lexical est celui des collocations du lexique scientifique transdisciplinaire.

**KEYWORDS** : collocation, modeling of collocations, syntactic properties, corpus linguistics

### **ABSTRACT**

Modeling the syntactic properties of collocations is not easy. This work focuses on the observation of syntactic constructions with the intent to understand how they work and propose a mechanism for encoding of these elements. The ultimate goal of this project is to provide the information necessary to create an online collocation dictionary. To that end, we used corpus linguistics methods. The lexical field is the transdisciplinary scientific lexicon of collocations.

