Université de Montréal

Étude comparative des combinaisons lexicales dans deux domaines de spécialité : collocations lexicales et collocations conceptuelles en aéronautique et en philosophie

par

Claudine Bertrand

Département de linguistique et traduction

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès arts (M.A.)
en traduction

Novembre 1998

© Claudine Bertrand, 1998

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Étude comparative des combinaisons lexicales dans deux domaines de spécialité : collocations lexicales et collocations conceptuelles en aéronautique et en philosophie

présenté par :

Claudine Bertrand

a été évalué par un jury composé des personnes suivantes :

Mémoire accepté le :

SOMMAIRE

Notre étude porte sur la combinatoire lexicale en langue de spécialité. Nous examinerons, plus précisément, les usages lexicaux de deux domaines spécifiques, l'aéronautique et la philosophie. Notre étude se penche particulièrement sur la distinction entre «collocation lexicale», combinaison dans laquelle un cooccurrent ne s'emploie qu'avec un seul terme, et «collocation conceptuelle», combinaison dans laquelle un cooccurrent s'emploie avec l'un ou l'autre terme d'une classe d'unités terminologiques, au sein des combinaisons lexicales spécialisées.

Dans une première étape, nous présentons au chapitre «État de la question» une analyse des points de vue de chercheurs en langue générale d'abord, puis en langue de spécialité. Nous terminons ce chapitre en mettant l'accent sur la distinction collocation lexicale versus collocation conceptuelle en langue de spécialité.

Le deuxième chapitre présente la méthodologie que nous avons adoptée pour mener à bien la présente recherche. Dans un premier temps, nous avons relevé les différents types de combinaisons par catégories (terme + verbe, verbe + terme, terme + adjectif, adjectif + terme, terme + préposition + nom, nom + préposition + terme) dans chacun des deux domaines. Ensuite, nous avons distingué les collocations lexicales des collocations conceptuelles, en examinant

dans quelle mesure les combinaisons lexicales pouvaient être généralisées à des ensembles de termes appartenant à une même classe conceptuelle.

Au chapitre des résultats, nous constatons que les collocations conceptuelles sont très productives en langue de spécialité, dans les deux domaines étudiés, mais que l'on y trouve également, dans une moindre mesure, des collocations lexicales.

TABLE DES MATIÈRES

SOMMAIRE iii

TABLE DES MATIÈRES v

LISTE DES TABLEAUX ix

LISTE DES FIGURES x

LISTE DES SIGLES ET ABRÉVIATIONS xii

REMERCIEMENTS xv

INTRODUCTION 1

CHAPITRE PREMIER : ÉTAT DE LA QUESTION 8

1.1 POINT DE VUE DE LA LANGUE GÉNÉRALE 13

1.1.1 Haussmann (1979) *13*

1.1.2 Benson et al. *(1986)* *15*

1.1.3 Mel'...uk et al. *(1995)* *16*

1.2 POINT DE VUE DE LA LANGUE SPÉCIALISÉE 24

1.2.1 Cohen (1983, 1986) *25*

1.2.2 Laporte (1996, 1997) *27*

1.2.3 Pavel (1993, 1994) *29*

1.2.4 Martin (1992) *31*

1.2.5 Heid (1992, 1994); Heid et Freibott (1991) *32*

1.3 COLLOCATION LEXICALE VS COLLOCATION CONCEPTUELLE 34

1.3.1 Langue générale *35*

1.3.1.1 Mel' uk et Wanner (1996) 35

1.3.2 Langue spécialisée *38*

1.3.2.1 Heid (1992, 1994) 38

1.3.2.2 Heid et Freibott (1991) 40

1.3.2.3 Meyer et Mackintosh (1994, 1996) 41

1.3.2.4 L'Homme (1993, 1995, 1998); L'Homme et Gemme (1997) 41

CHAPITRE II : MÉTHODOLOGIE 47

2 Processus de recherche 48

2.1 CONSTITUTION DU CORPUS 49

2.1.1 Choix des domaines étudiés 49

2.1.2 Ouvrages portant sur le domaine de l'aéronautique 49

2.1.3 Ouvrages portant sur le domaine de la philosophie 50

2.1.4 Choix de la langue 50

2.1.5 Échantillonnage 51

2.1.6 Numérisation 51

2.1.6.1 Lecture optique des documents 51

2.1.6.2 Correction des documents 53

2.2 SÉLECTION DES TERMES CLÉS 54

2.3 RECHERCHE DE CONTEXTES 56

2.3.1 Présentation du logiciel Naturel Pro^MC 56

2.4 SÉLECTION DES COMBINAISONS LEXICALES SPÉCIALISÉES (CLS) 59

2.4.1 Définition 59

2.4.2 Terme complexe et CLS 60

2.4.3 Formes de CLS retenues 62

2.5 CONSTITUTION D'UNE BASE DE DONNÉES À L'AIDE DU LOGICIEL EXCEL® 63

2.5.1 Présentation du logiciel Excel® 63

2.5.2 Constitution de la base de données 64

2.5.3 CLS — domaine de l'aéronautique 68

2.5.4 CLS — domaine de la philosophie 69

2.5.5 Aéronautique vs philosophie 70

2.6 INVENTAIRE DES COOCCURRENTS 70

2.7 DEUXIÈME RECHERCHE À PARTIR DES COOCCURRENTS 74

2.8 REGROUPEMENT DES TERMES PAR CLASSES CONCEPTUELLES 79

2.8.1 Détermination des classes conceptuelles 80

2.8.2 Élaboration des classes 81

2.8.3 Regroupement.....83

CHAPITRE III : RÉSULTATS..... 85

3 Considérations générales..... 85

3.1 RÉPARTITION DES COLLOCATIONS — DOMAINE DE L'AÉRONAUTIQUE.....86

3.1.1 Collocations conceptuelles..... 89

3.1.2 Collocations lexicales.....92

3.2 RÉPARTITION DES COLLOCATIONS — DOMAINE DE LA PHILOSOPHIE.....95

3.2.1 Collocations conceptuelles..... 99

3.2.2 Collocations lexicales.....100

3.3 COMPARAISONS ENTRE LES DEUX DOMAINES 103

CONCLUSION 105

BIBLIOGRAPHIE 108

ANNEXE 1 : OUVRAGES PORTANT SUR L'AÉRONAUTIQUE xvi

ANNEXE 2 : OUVRAGES PORTANT SUR LA PHILOSOPHIE xvii

ANNEXE 3 : LISTE DE COOCCURRENTS DU DOMAINE DE L'AÉRONAUTIQUE.....xviii

ANNEXE 4 : LISTE DE COOCCURRENTS DU DOMAINE DE LA PHILOSOPHIE xix

ANNEXE 5 : DEUXIÈME RECHERCHE DE COOCCURRENTS xx

ANNEXE 6 : RÉSULTATS — AÉRONAUTIQUE xxi

ANNEXE 7 : RÉSULTATS — PHILOSOPHIE xxii

ANNEXE 8 : REGROUPEMENT PAR CLASSES CONCEPTUELLES.....xxiii

ANNEXE 9 : EXTRAIT DE NOTRE BASE DE DONNÉES (CORPUS) xxiv

ANNEXE 10 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES VERBALE.....xxv

ANNEXE 11 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES NOMINALE..... xxvi

ANNEXE 12 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES ADJECTIVALE xxvii

LISTE DES TABLEAUX

Tableau I : Termes clés sélectionnés 55

Tableau II : Catégories de combinaisons lexicales spécialisées..... 62

Tableau III : Résultats des combinaisons obtenues — aéronautique 69

Tableau IV : Résultats des combinaisons obtenues — philosophie..... 70

Tableau V : Cooccurrents sélectionnés dans la base Aéronautique..... 72

Tableau VI : Cooccurrents sélectionnés dans la base Philosophie..... 73

Tableau VII : Deuxième recherche de termes — *autoriser* 74

Tableau VIII : Deuxième recherche de termes — *principe*..... 77

Tableau IX : Regroupement à partir du verbe *autoriser* 83

Tableau X : CC vs CL en fonction des classes conceptuelles (aéronautique) 87

Tableau XI : CC vs CL en fonction des catégories grammaticales (aéronautique) 88

Tableau XII : Fréquences observées — *exploiter* 91

Tableau XIII : Fréquences observées — *maîtriser* 91

Tableau XIV : CC vs CL en fonction des classes conceptuelles (philosophie). 96

Tableau XV : CC vs CL en fonction des catégories grammaticales (philosophie) 98

LISTE DES FIGURES

Figure 1 : Exemple de collocation conceptuelle 4

Figure 2 : Rapports entre lexèmes 18

Figure 3 : Fonction lexicale 20

Figure 4 : Vocable «chagrin» 21

Figure 5 : Entrée du *Lexique de cooccurrents — Bourse et conjoncture économique* (Cohen 1986 : 54) 26

Figure 6 : Entrée *dose* 28

Figure 7 : Noms dénotant une émotion et leurs composantes sémantiques 36

Figure 8 : Valeurs des composantes sémantiques 37

Figure 9 : Regroupement des noms dénotant une augmentation ou une diminution 39

Figure 10 : Combinaison ABSORB + USED WATERS 43

Figure 11 : Entrée d'une base de données terminologique incluant un enregistrement verbal 44

Figure 12 : Diagramme de cheminement 48

Figure 13 : Exemple de texte numérisé 53

Figure 14 : Écran de saisie des quatre fenêtres d'affichage dynamique proposées par Naturel Pro^{MC} 58

Figure 15 : Exemple d'enregistrement de la base Aéronautique 64

Figure 16 : Exemple d'enregistrement verbal 66

Figure 17 : Exemple d'enregistrement nominal 66

Figure 18 : Exemple d'enregistrement adjectival 67

Figure 19 : Listes partielles de cooccurrents 67

Figure 20 : Distinction entre les différents sens de *autoriser* 76

Figure 21 : Distinction entre les différents sens de *principe* 78

LISTE DES SIGLES ET ABRÉVIATIONS

A :	adjectif
AC :	activité
ACT :	actant
ADV :	adverbe
AM :	activité morale
ANTÉ :	antéposition
ART DÉF :	article défini
ART IND :	article indéfini
ATC :	contrôle de la circulation aérienne
AV :	appareil volant
CC :	collocation conceptuelle
CL :	collocation lexicale
CLS :	combinaison lexicale spécialisée
COI :	complément d'objet indirect
CTC :	contexte du terme clé
CTX :	contexte
DI :	disposition
EA :	entité animée
EB :	entité abstraite
EI :	entité immatérielle
EM :	entité matérielle

EP :	entité physique
FL :	fonction lexicale
FSS :	station d'information de vol
IFR :	règles de vol aux instruments
IT :	installations propres au trafic aérien
LG :	langue générale
LI :	lieu
LS :	langue spécialisée
N :	nom
NPTC :	nom + préposition + terme clé
OBJ :	objet
POST :	postposition
PR :	principe
PRÉP :	préposition
SUJ :	sujet
T :	terme
TCA :	terme clé + adjectif
TCD :	terme clé de départ
TCOD :	terme clé + complément d'objet direct
TCPN :	terme clé + préposition + nom
TCV :	terme clé + verbe
UM :	unité de mesure
V :	verbe

VC :	voie de circulation
VFR :	règles de vol à vue
VTC :	verbe + terme clé

REMERCIEMENTS

Nous tenons à remercier chaleureusement notre directrice de recherche, madame Marie-Claude L'Homme, professeure au département de linguistique et de traduction.

Par ses encouragements constants, sa patience inébranlable, ses conseils toujours pertinents et sa grande disponibilité, elle nous a aidée à franchir chacune des étapes qui ont mené à la réalisation de ce mémoire.

Nous tenons également à souligner notre reconnaissance à mesdames Aline Francoeur, Chantal Litalien et Brigitte Morel, qui ont généreusement accepté de relire notre travail. Nous remercions aussi monsieur Marcel Paquet pour ses précieux conseils en matière d'informatique.

Enfin, nous tenons à exprimer notre plus vive gratitude à Jean-François et Simon Provencher pour leur appui inconditionnel.

INTRODUCTION

Selon Sager (1990), le lexique d'une langue spécialisée, du point de vue de la terminologie, consiste en de nombreux sous-systèmes qui représentent la structure du savoir dans un champ ou une discipline donnés. L'auteur souligne que ces structures ne constituent pas des entités absolues, mais qu'elles sont le reflet de l'état du savoir d'un individu ou d'un groupe de spécialistes. Il se dégage donc une caractéristique importante de ces systèmes conceptuels en ceci qu'il s'agit d'entités mouvantes en constante évolution, principalement dans les secteurs de la recherche et du développement de la science et de la technologie, d'où, notamment, la difficulté d'une dénomination stable.

Or, lorsqu'il s'agit de groupements de mots liés à un usage établi dans une discipline particulière, cette mouvance terminologique constitue à tout le moins un obstacle à leur délimitation. Pourtant, soulignons qu'un utilisateur de la langue de spécialité (qu'il s'agisse d'un communicateur, d'un scientifique ou d'un professionnel langagier) doit connaître et «re-connaître» les conventions liées à la configuration du savoir qui sous-tend le discours en question. De là l'importance de considérer l'environnement phrastique des termes.

Ainsi, lorsqu'il s'agit de traduire les textes et techniques relevant de domaines de spécialité, l'un des problèmes qui se posent, au-delà du souci de l'exactitude du contenu, c'est de trouver la bonne tournure, la façon de dire propre

au domaine concerné. Bien sûr, les traducteurs disposent d'une panoplie d'outils (banques de terminologie, dictionnaires spécialisés, lexiques, etc.), mais ces derniers fournissent des équivalences lexicales sans pour autant situer les unités terminologiques dans leur environnement phrastique. D'emblée, nous pouvons donc affirmer que le professionnel de la langue a besoin d'outils terminologiques qui soient à la fois conceptuels (équivalence notion — terme) et contextuels (comportement de ce terme par rapport aux autres éléments constitutifs de la phrase).

Le présent mémoire a pour but de mieux comprendre les groupements lexicaux privilégiés de la langue spécialisée, que nous appellerons «combinaisons lexicales spécialisées» (CLS), c'est-à-dire des groupes composés de deux lexèmes utilisés selon une certaine convention au sein d'une communauté de spécialistes d'un domaine donné et dont l'un des constituants est un «terme» (*piloter un* ***aéronef****, conscience de* ***soi****,* ***moi*** *spirituel*).

Certaines particularités de la langue se révèlent difficiles à expliquer par le seul recours aux règles syntaxiques ou sémantiques. En effet, un apprenant de la langue, par exemple, ne fera probablement aucune différence entre les expressions *conduire une automobile* et *conduire un aéronef*, puisque, dans les deux cas, i) elles se soumettent à la définition du terme *conduire* donnée dans Le nouveau petit Robert (1995 : 435), soit «diriger ou manoeuvrer un engin», et ii) leur construction ne fait aucunement entorse à la syntaxe française.

Or, si l'usage sanctionne l'utilisation de la première expression, il en est autrement pour ce qui est de la deuxième : dans le milieu aéronautique, par suite d'une convention établie au sein des spécialistes, on exprime la notion de «diriger un aéronef» par l'expression *piloter un aéronef,* le terme *piloter* donnant lieu à d'autres groupements privilégiés tels que *pilote d'aéronef* — plutôt que *conducteur d'aéronef,* ou encore *pilotage automatique* — plutôt que *conduite automatique.*

Ces phénomènes revêtent une importance certaine; pourtant, leurs aspects fondamentaux n'ont pas fait l'objet d'études approfondies en terminologie. Voici quelques-unes des questions qui ont été soulevées :

- Est-il possible de déceler des tendances au niveau du comportement syntagmatique des combinaisons lexicales spécialisées?

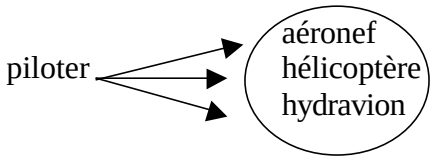
- Étant donné les difficultés de découpage des combinaisons lexicales, à partir de quel moment le passage se fait-il entre termes complexes et autres formes de combinaisons de mots?

- Ces difficultés sont-elles plus nombreuses dans le domaine des sciences humaines où la commutabilité des termes semble beaucoup plus grande (du fait de l'usage prédominant de mots de la langue générale) que dans le domaine des sciences dites exactes marqué par un foisonnement d'unités terminologiques complexes et où la commutabilité est restreinte?

- Peut-on envisager de répertorier les combinaisons lexicales spécialisées par domaine, de les typologiser selon leurs structures?

Nous aborderons certaines de ces questions indirectement, mais notre mémoire se penche plus particulièrement sur le rapport entre collocation lexicale (CL) et collocation conceptuelle (CC). De façon sommaire, nous exprimons la différence entre ces deux types de CLS de la manière suivante : en admettant que la CLS se compose de deux unités 1) une base, qui est un terme de nature nominale, et 2) un cooccurrent, on entend par collocation lexicale une combinaison dans laquelle un cooccurrent ne s'emploie qu'avec un seul terme, et par collocation conceptuelle une combinaison dans laquelle un cooccurrent s'emploie avec une classe d'unités terminologiques. Dans la figure ci-dessous, nous présentons un exemple de collocation conceptuelle dans laquelle le cooccurrent *piloter* peut s'employer avec une classe d'unités terminologiques qui comprend notamment les termes *aéronef*, *hélicoptère* et *hydravion*.

Figure 1 : Exemple de collocation conceptuelle



À l'aide d'un corpus, nous vérifierons ce rapport entre CL et CC en sélectionnant d'abord des combinaisons lexicales spécialisées (CLS) à partir de termes clés de départ. Puis, cette fois à partir des cooccurrents retenus, nous sélectionnerons d'autres CLS. Enfin, nous établirons des classes conceptuelles qui nous permettront de regrouper les CC, et de mesurer, en distinguant les CC des CL, combien de CL et de CC nous aurons obtenues dans les deux domaines concernés.

Notre travail pourrait mener à l'élaboration d'une méthode de confection de lexiques de cooccurrents, informatisés ou non. Nous croyons que l'étude des combinaisons lexicales permettra également d'enrichir les banques de données terminologiques existantes par l'intégration de l'information relative à la combinatoire terminologique, et ainsi, de combler un besoin de plus en plus pressant chez les professionnels de la langue.

Par ailleurs, nous croyons que notre étude sera utile à d'autres points de vue, ne serait-ce que pour confirmer l'importance de saisir les contraintes qui s'exercent sur l'association des éléments linguistiques en unités organisées. Nous pensons ici à la recherche en linguistique computationnelle, qui s'intéresse particulièrement aux rapports existant entre les termes et les unités lexicales avec lesquelles ces termes se combinent.

Notre mémoire comprend trois volets :

1. Le premier chapitre présente un survol des travaux effectués sur les collocations dans la langue générale ainsi que de quelques études menées dans des domaines spécialisés. Nous y faisons état des différentes caractéristiques des combinaisons lexicales (ci-après appelées «collocations» en langue générale et «combinaisons lexicales spécialisées» en langue de spécialité). Nous exposons diverses argumentations et méthodes utilisées par les chercheurs, pour en arriver à examiner une problématique qui se dessine graduellement en terminologie, c'est-à-dire celle des collocations lexicales par rapport aux collocations conceptuelles.

2. Le deuxième chapitre expose en détail la méthodologie de notre recherche. À partir d'un dépouillement assisté par ordinateur (et à l'aide de différents logiciels tels que NaturelPro^{MC} et Excel®), nous avons constitué un corpus composé de textes relatifs à l'aéronautique ainsi qu'à la philosophie. Nous avons sélectionné des termes clés de départ, puis répertorié, dans trois principales catégories (verbale, nominale et adjectivale) des cooccurrents. Une fois cette étape franchie, nous avons fait l'inventaire des cooccurrents relevés, et, à partir de ces derniers, nous avons répertorié d'autres termes, que nous avons tenté de regrouper en classes conceptuelles.

3. Le troisième chapitre livre les résultats que nous avons obtenus. En répartissant les CLS en collocations lexicales et en collocations conceptuelles en fonction des caractéristiques retenues au cours de notre recherche, nous dégageons certaines observations quant au comportement des CLS dans les deux domaines à l'étude.

CHAPITRE PREMIER : ÉTAT DE LA QUESTION

Les chercheurs et autres spécialistes de la langue (terminologues, linguistes, lexicographes et professionnels langagiers) ont recours à diverses dénominations pour désigner les groupements privilégiés : *collocation* (Mel' uk *et al.* : 1984, 1988, 1992, 1995), *phraséologisme* (Pavel : 1993, 1994), *cooccurrents* (Cohen : 1983, 1986), *combinaison lexicale* (L'Homme : 1993, 1995, 1998). Notre propos n'est pas de nous pencher sur ces variantes, aussi, pour les besoins de notre recherche, avons-nous adopté la terminologie proposée par L'Homme (1998), à savoir CLS (*combinaison lexicale spécialisée*) pour les groupements de la langue spécialisée, et *collocation* pour ceux de la langue générale. Ces choix ont pour unique but de clarifier l'exposé qui suit.

Il nous paraît important de distinguer, dans une première étape, les points de vue des différents auteurs en fonction de la distinction langue générale/langue spécialisée. Soulignons toutefois qu'en raison des limites imposées par ce mémoire, nous avons dû restreindre le nombre d'auteurs à l'étude, pour ne sélectionner que ceux qui semblaient les plus pertinents à notre recherche.

Nous examinerons, dans un premier temps (point 1.1), comment les différents auteurs définissent la collocation en langue générale et, dans un deuxième temps (point 1.2), en langue de spécialité, pour en venir, et de manière plus spécifique (point 1.3), à la notion de «collocation lexicale» vs «collocation

conceptuelle» telle qu'elle est traitée par certains chercheurs. Pour ce faire, nous avons examiné les points suivants :

- cadre dans lequel l'auteur s'inscrit;
- champ de spécialisation;
- argumentation;
- méthode utilisée.

Nous avons tenté, dans la mesure du possible, de départager les observations faites sur les collocations de celles qui ont été formulées sur les CLS. Nous verrons toutefois que les propos de certains auteurs touchent à la fois la langue générale et la langue spécialisée, certains ne faisant pas de distinction entre les groupements.

Pour mieux distinguer les groupements de mots abordés dans le présent mémoire, nous avons dressé une liste de définitions des principales notions. Dans les pages qui suivent, nous prêtons les sens suivants aux termes utilisés :

Collocation : On appelle *collocation* l'association imposée par l'usage de deux lexèmes qui entretiennent une relation polaire. Le premier lexème — *base* (Haussmann : 1979) ou *mot clé* (Mel' uk *et al.* : 1995) se combine avec un second — *collocatif* (Haussmann : 1979) ou *valeur de la fonction lexicale* (Mel' uk *et al.* : 1995).

Exemple | *joie débordante* (Clas 1994 : 577)

joie = base (ou mot clé)

débordante = collocatif (ou valeur de la fonction lexicale)

Combinaison lexicale spécialisée (CLS) : La CLS présente un caractère conventionnel au sein d'une communauté de spécialistes. Elle se compose de deux lexèmes, dont une unité terminologique, généralement de nature nominale — *base* ou *terme noyau* (Pavel : 1994) combinée à un verbe, un nom ou un adjectif — *cooccurrent* (Pavel : 1994).

Exemple | *aéronef civil* (notre corpus)

aéronef = base (ou terme noyau)

civil = cooccurrent

Cooccurrent : En terminologie, le cooccurrent est un élément qui est adjoint au terme (la base) dans une association dite non libre et régie par un usage établi.[1]

[1] Notons que l'on prête parfois au terme *cooccurrence* l'extension suivante en linguistique : «Si on appelle occurrence d'un élément linguistique *x* le fait que *x* figure dans une phrase donnée, les éléments qui figurent avec lui dans cette phrase sont ses *cooccurrences.* On dit que les éléments B, C et D sont *cooccurrents* d'un autre élément A quand A figure avec ces éléments, chacun dans une position déterminée, pour produire un énoncé» (Dubois *et al.* : 1994).

Exemple | *prix exorbitant* (Cohen 1983 : 6)

prix = base

exorbitant = cooccurrent

Expression idiomatique : On appelle *expression idiomatique* toute forme grammaticale dont le sens ne peut être déduit de sa structure en morphèmes et qui n'entre pas dans la constitution d'une forme plus large (Dubois *et al.* : 1994).

Exemple | *arriver comme un cheveu sur la soupe* (Mel' uk *et al.* : 1995)

Concept : Le concept, souvent appelé *notion* en terminologie, est une représentation mentale d'une réalité concrète ou abstraite ou d'une classe d'objets, exprimée par un terme ou par un symbole (ministère du Secrétariat d'État du Canada : 1996).

Phraséologisme : Sous-système linguistique qui utilise une terminologie et d'autres moyens linguistiques et qui vise la non-ambiguï té de la communication dans un domaine particulier (Norme ISO 1087).[2]

Exemple | *À la discrétion du pilote* (AIP Canada : 1995)

[2] Cette définition est celle qui est retenue par la terminologie. En linguistique, *phraséologisme* est souvent utilisé comme équivalent de *expression idiomatique*.

Phrasème : Un *phrasème* est une locution prise dans une seule acception bien déterminée et munie de tous les renseignements qui spécifient totalement son comportement dans un texte (Mel' uk *et al.* : 1995).[3]

Terme (unité terminologique) : En terminologie, le *terme* ou *unité terminologique* est l'unité signifiante constituée d'un mot (*terme simple*) ou de plusieurs mots (*terme complexe*), qui désigne une notion de façon univoque à l'intérieur d'un domaine (définition de l'Office de la langue française du Québec) (Dubois *et al.* : 1994).

Exemples	terme simple : *aéronef* (AIP Canada : 1995)
	terme complexe : *espace aérien* (AIP Canada : 1995)

[3] Nous verrons que Mel' uk *et al.* distinguent trois catégories de *phrasèmes*, soit les «phrasèmes complets», les «semi-phrasèmes» et les «quasi-phrasèmes».

1.1 Point de vue de la langue générale

Chez les spécialistes qui ont traité de la combinatoire lexicale en langue générale, nous avons retenu Haussmann (1979), Benson *et al.* (1986) et Mel' uk *et al.* (1995).

1.1.1 Haussmann (1979)

Les travaux de Hausmann s'inscrivent dans le cadre de la lexicographie. En effet, l'auteur s'intéresse particulièrement aux implications dictionnairiques des collocations, c'est-à-dire le recensement des collocations et l'encodage de celles-ci dans des ouvrages de référence. Il s'est principalement attaqué aux problèmes liés à la production de textes en langue étrangère, visant directement les professionnels langagiers tels que les traducteurs. Haussmann accorde donc une grande importance aux collocations, qu'il qualifie de «produits semi-finis de la langue», leur accordant une place prépondérante dans le processus rédactionnel. Haussmann insiste sur le fait que l'apprentissage des collocations favorise la production de textes plus précis au point de vue sémantique, et leur assure l'idiomaticité souhaitée et nécessaire.

Pour Haussmann, la collocation se compose de deux éléments appelés *base* et *collocatif*, et leur relation est qualifiée de «polaire» : la base conserve son sens

(une *sémantique autonome*), contrairement au collocatif (*sémantiquement dépendant*), qui ne s'interprète qu'en fonction de la base. Il existe donc une contrainte lexicale pour ce deuxième élément. Haussmann ajoute qu'il existe une *affinité* entre les deux éléments et que leur combinaison est imprévisible, d'où la nécessité de mémoriser les collocations (citons quelques exemples donnés par l'auteur : *refus catégorique, preuve irréfutable, ferme résolution*).

À la suite de ces analyses, Haussmann a classé les collocations en se basant sur une typologie catégorielle des éléments (nom + nom — *pincée de sel*; verbe + nom — *prononcer un discours*; nom + adjectif — *marque distinctive*; adverbe + adjectif — *vachement bon*; verbe + adverbe — *boire goulûment*[4]).

La collocation telle qu'elle est décrite par Haussmann se caractérise de la manière suivante : i) deux lexèmes qui entretiennent une relation polaire; ii) ceux-ci formant une combinaison lexicalement contrainte; et iii) le choix des lexèmes pouvant se combiner avec une base donnée est imprévisible.

---

[4] Exemples tirés de Clas (1994 : 578).

1.1.2 Benson *et al*. (1986)

Benson *et al*. ont publié un dictionnaire (unilingue anglais)[5] qui rend compte de la combinatoire des entrées. Dans cet ouvrage, les auteurs définissent la collocation comme étant une combinaison de mots récurrente (*recurrent word combination*) et semi-figée. Benson *et al.* répartissent les collocations en deux groupes, soit les collocations grammaticales et les collocations lexicales :

- les collocations grammaticales consistent en un mot dominant — nom, adjectif ou verbe — suivi d'une préposition (ex. : *admiration for*) ou d'une structure grammaticale telle qu'une proposition infinitive ou une proposition relative (ex. : *compulsion to do it; we reach an agreement that she would represent us in court*)[6];

- les collocations lexicales (ex. : *world championship, formidable challenge*), contrairement aux collocations grammaticales, ne comportent pas de préposition, ni de propositions infinitive ou relative; ce sont des combinaisons de lexèmes. Elles se construisent essentiellement selon la typologie suivante : verbe + nom (*set an alarm*); adjectif + nom (*strong tea*); nom + verbe (*alarm go off*); adverbe + verbe (*deeply absorbed*); verbe + adverbe (*affect*

---

[5] *The BBI Combinatory Dictionary of English; A Guide to Word Combinations* (1986).
[6] Il est intéressant de noter que, parmi les auteurs étudiés, Benson *et al*. sont les seuls à traiter ces groupes comme des collocations. Ajoutons que ce sont les collocations du 2e groupe, à savoir les collocations lexicales, qui nous intéressent le plus ici.

deeply). De plus, à l'instar de Haussmann (1979), Benson *et al.* reconnaissent que les collocations de ce type sont arbitraires et imprévisibles, et qu'il est donc nécessaire pour les non-natifs de la langue de les apprendre.

Pour Benson *et al*, les collocations lexicales sont : i) des combinaisons de mots récurrentes et semi-figées, ii) elles sont imprévisibles.

1.1.3 Mel' uk *et al*. (1995)

Mel' uk *et al.* présentent un modèle (lexicologie explicative et combinatoire)[7] qui permet de décrire de façon systématique les unités lexicales de la langue générale. Les expressions lexicales ainsi décrites peuvent être plus ou moins figées, elles peuvent être des expressions idiomatiques ou non, et elles présentent un caractère conventionnel. C'est pourquoi tout locuteur non natif de la langue doit les apprendre s'il désire communiquer de manière juste et utiliser les «bonnes tournures» propres à une langue donnée.

Mel' uk *et al.* situent la collocation parmi d'autres associations de mots regroupées sous le générique *phrasème.* Les phrasèmes sont classés en trois catégories : les phrasèmes complets, les quasi-phrasèmes et les semi-phrasèmes.

---

[7] À noter que les modèles lexicaux s'insèrent dans un modèle linguistique plus vaste appelé «sens-texte».

Les phrasèmes complets sont des expressions dont le sens n'est pas compositionnel, c'est-à-dire que le sens global ne peut être déduit à partir du sens des éléments lexicaux qui forment l'expression. Il s'agit d'expressions idiomatiques comme *pont aérien* (liaison régulière par avion entre deux points)[8].

Les quasi-phrasèmes sont des locutions qui ne peuvent pas être complètement décrites en fonction d'au moins un de leurs constituants, par exemple, *centre commercial*[9].

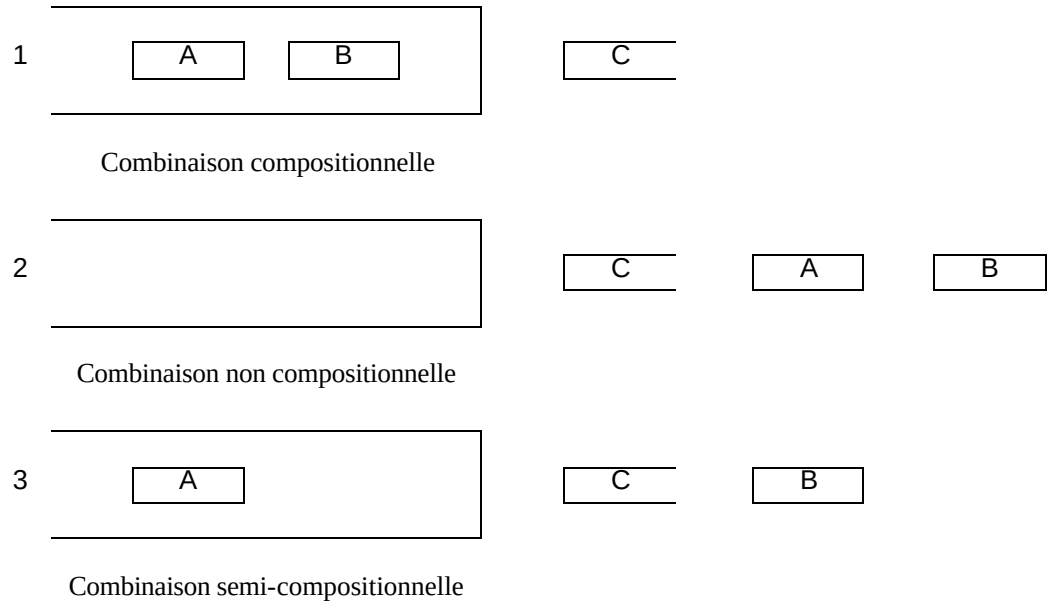
Dans les semi-phrasèmes, l'un des éléments — le mot clé — garde son sens propre, tandis que l'autre (voir la figure 2 ci-après) est lexicalement contraint. Par exemple, *piquer un somme* inclut le sens de *somme* mais pas celui de *piquer*[10]. Les collocations font partie de ce troisième groupe.

Un nouvel examen des rapports entre lexèmes permet de situer les collocations. La figure 2 ci-dessous reprend un schéma que Heid (1994) a retenu lors d'une rencontre de travail avec Mel' uk en 1989 :

---

[8] Exemple tiré de Mel' uk *et al*. (1995 : 16).
[9] Exemple tiré de Mel' uk *et al*. (1995 : 16).
[10] Exemple tiré de Mel' uk *et al*. (1995 : 16).

Figure 2 : Rapports entre lexèmes



Soit la combinaison A + B = C.

Dans l'exemple 1 : A = col

B = bleu

La résultante C, *col bleu* («j'aime le *col bleu* de cette blouse»), est la somme des deux éléments de la relation, et elle est qualifiée de combinaison compositionnelle.

Dans l'exemple 2 : A = pont

B = aérien

La résultante C, *pont aérien*, constitue une expression idiomatique dans laquelle on ne reconnaît ni le sens de A, ni le sens de B. La combinaison est alors dite non compositionnelle (phrasème complet).

Dans l'exemple 3 : A = somme

B = piquer

La résultante C, *piquer un somme*, conserve le sens de A, mais pas celui de B, et la relation est semi -compositionnelle (semi-phrasème).

Pour Mel' uk *et al.*, les collocations sont des groupes «... lexicalement contraints, c'est-à-dire dont la combinatoire n'est déterminée ni par leur sémantisme ni par leurs propriétés syntaxiques» (Mel' uk *et al*. 1995 : 126). Par exemple, pourquoi, se demandent les auteurs, peut-on dire *avoir recours à* et non **faire recours à*? (Mel' uk *et al*. 1995 : 126). Aucune règle sémantique ni syntaxique ne justifie l'emploi d'une expression de préférence à l'autre. Autrement dit, les collocations ne sont pas prévisibles.

Mel' uk *et al*. classent les collocations en fonction du rapport sémantique existant entre leurs composantes et les présentent d'après ce qu'ils nomment des fonctions lexicales (FL), lesquelles constituent une «constante sémantique». On peut comparer cette fonction lexicale à une formule mathématique : f (x) = y, dans laquelle «f» constitue la fonction exprimée, «x» l'argument (le mot clé) et «y» la valeur, c'est-à-dire la résultante. Par exemple, on peut exprimer la constante

sémantique d'intensification de multiples façons. Mel' uk *et al*. ont attribué à cette constante sémantique le nom de fonction Magn. Voici quelques exemples de cette fonction :

Figure 3 : Fonction lexicale

1. Magn (refuser) =	absolument, catégoriquement
2. Magn (peur) =	bleue
Constante sémantique =	intensification (sens 'très')
Fonction lexicale =	Magn
Mot clé =	1. refuser
	2. peur
Valeur =	1. absolument, catégoriquement
	2. bleue[11]

Nous reproduisons ci-dessous une partie de l'entrée du vocable «chagrin» traité dans le *Dictionnaire explicatif et combinatoire du français contemporain* (Mel' uk *et al*. 1992 : 164).

[11] Notons que certaines FL sont utilisées à d'autres fins, entre autres pour rendre compte de relations paradigmatiques telles que la synonymie ou l'antonymie :

Exemple :	Syn (refuser) =	décliner
Dans cet exemple :	la fonction lexicale =	Syn
	le mot clé =	refuser
	la valeur =	décliner

Figure 4 : Vocable «chagrin»

CHAGRIN

I.1. Forte émotion désagréable ... [*le chagrin de ma soeur à la suite de son échec*]

2. Fait Y qui cause le chagrin I.1. ... [*Son grand chagrin fut la mort de son chat*]

I.1. nom, masc. *Chagrin de X à cause de Y*= Forte émotion désagréable de X qui est causée par un fait Y contraire aux profonds désirs de X et devant lequel X se croit impuissant, et qui cause chez X un état de tristesse et l'envie 1 de ne rien faire; cette émotion est celle qu'on a normalement dans de pareilles situations.

Régime

1 = X	2 = Y
1. *de* N 2. A_{poss} 3. A	1. *à cause de* N 2. *à la suite de* N 3. *devant* N 4. *de* V_{inf}

C_1 : *le chagrin de Simon, son chagrin, le chagrin paternel*

C_2 : *le chagrin à cause d'une séparation (à la suite d'un échec, devant tant de misère, de devoir partir*)

$C_1 + C_2$: *le chagrin de Simon (son chagrin) à cause de sa séparation (à la suite de son échec, devant cette injustice, de devoir partir); leur chagrin à*

cause de la perte d'un homme si juste (à la suite de sa démission, devant cette indifférence, d'avoir perdu cet homme)[12]

Fonctions lexicales

AntiMagn : petit (*prépos)*

$Magn^{temp}$: long (*prépos*), durable, éternel, perpétuel

IncepPredPlus : s'accroître, augmenter, grandir

IncepPredMinus : diminuer, faiblir, se dissiper [*Avec le temps, le chagrin se dissipe*]

CausPredMinus : diminuer [ART ~]; apaiser, calmer, soulager [ART ~] [*J'essayais de calmer le chagrin de Simon*]

$Oper_1$: avoir, éprouver, ressentir [du ~]

Magn + $Oper_1$: être tout(e) [à $ART_{déf}$ ~] [*Marie était toute à son chagrin*], avoir le coeur gonflé (plein, rongé) [de ~][13]

Ces fonctions lexicales sont développées dans le modèle «sens-texte» qui, soulignons-le, laisse une large place à la combinatoire lexicale. Le métalangage utilisé dans le modèle «sens-texte» constitue une excellente plateforme pour l'étude des combinaisons lexicales. En effet, le formalisme des fonctions lexicales

---

[12] N = nom; Aposs = adjectif possessif; C0 = mot clé; Vinf = verbe infinitif; prépos = préposé; ART = article; $ART_{déf}$ = article défini.

[13] Pour connaître la liste des fonctions lexicales standard et non standard établies par Mel' uk et son équipe, consulter un des volumes du *Dictionnaire explicatif et combinatoire du français contemporain* (Mel' uk *et al*, 1984, 1988, 1992).

offre l'avantage d'être systématique et logique. C'est d'ailleurs pourquoi plusieurs chercheurs ont eu recours à ce formalisme ou à un mécanisme comparable, notamment Cohen (1983), Fontenelle (1996), Heid (1992, 1994) et Heid et Freibott (1991) (nous y reviendrons plus loin).

Cependant, les FL, qui fonctionnent bien lorsqu'il s'agit de décrire les collocations de la langue générale, seraient, selon certains auteurs, moins maniables pour ce qui est des langues spécialisées, pour lesquelles on devrait recourir à des fonctions lexicales non standard :

> «...Les FL non standard sont surtout typiques pour des mots concrets, qui sont normalement très marqués culturellement ou techniquement : noms de nourriture [...], de transactions financières, de procédures médicales, d'activités culturelles, politiques ou religieuses, d'armes, de parties du corps, ... Ces fonctions sont extrêmement nombreuses dans le secteur des langues spécialisées (technologie, enseignement, droit, ...). Leur recensement et leur description lexicographique représentent une tâche fort ardue» (Mel' uk *et al*. 1995 : 151).

Nous verrons toutefois que certains auteurs ont eu recours à des mécanismes semblables aux FL pour rendre compte de combinaisons lexicales dans des domaines spécialisés.

Des travaux de Mel' uk *et al*., il ressort que les collocations sont : i) lexicalement contraintes; ii) arbitraires, donc, par voie de conséquence; iii)

imprévisibles; iv) semi-compositionnelles; et v) les liens sémantiques unissant un mot clé à son collocatif sont généralisables.

Nous retrouvons chez les auteurs que nous venons d'étudier plusieurs similitudes. Pour Haussmann (1979), Benson *et al*. (1986) ainsi que Mel' uk *et al*. (1995), la collocation se compose de deux éléments (appelés *base* et *collocatif* par Haussmann; *mot clé* et *valeur de la fonction lexicale* par Mel' uk *et al*.) et leur association est de nature conventionnelle et imprévisible. De plus, les éléments qui composent la collocation sont lexicalement contraints, et la collocation peut être semi-compositionnelle. Mel' uk *et al*. ont en outre découvert qu'il était possible de généraliser les liens sémantiques existant entre les éléments.

Au point 1.2 ci-après, nous examinons le point de vue de quelques autres chercheurs touchant plus particulièrement les groupements de mots de la langue spécialisée.

1.2 Point de vue de la langue spécialisée

Parmi les nombreux auteurs qui se sont intéressés aux combinaisons lexicales spécialisées, nous considérerons, dans l'ordre, Cohen (1983, 1986), Laporte (1996, 1997), Pavel (1994), Martin (1992), Heid (1992, 1994), Heid et Freibott (1991), Meyer et Mackintosh (1996), L'Homme (1993, 1995, 1998) et, enfin, L'Homme et Gemme (1997).

1.2.1 Cohen (1983, 1986)

Cohen a produit le *Lexique de cooccurrents — Bourse et conjoncture économique* (1986). Son travail s'inscrit donc en quelque sorte dans le cadre de la lexicographie relative à un champ de spécialisation bien précis, l'économie. Cohen distingue trois sortes d'associations de mots dans la langue : l'association libre, l'expression figée et le groupement usuel (Cohen 1983 : 5-6). C'est cette dernière association que Cohen appelle *cooccurrence lexicale* telle que décrite par Bally (1944). Elle met l'accent sur le caractère usuel des combinaisons lexicales définies comme des «groupements préférentiels», une «association de mots consacrés par l'usage», par exemple, *accélération rapide*, [l']*argent se déprécie*, *souscrire une action* (Cohen 1983 : 7).

Pour classer les associations de ce type, Cohen a eu recours à un mécanisme qui s'apparente au modèle des fonctions lexicales élaboré par Mel' uk *et al.* (1984, 1988, 1992, 1995) dont nous avons exposé les grandes lignes au point 1.1.3. L'auteure a établi sa classification en fonction des phases du cycle économique (début, croissance, déclin et fin) et les cooccurrents ont été regroupés par catégories grammaticales (nom, verbe/sujet, verbe/objet et adjectif).

En s'inspirant du modèle des fonctions lexicales et en s'appuyant sur une classification d'ordre conceptuel, Cohen a donc pu systématiser les relations sémantiques existant entre termes et cooccurrents dans le domaine qu'elle a choisi

d'étudier. La figure 5 ci-après présente une entrée du lexique de Cohen qui traite du terme *emprunt*.

Figure 5 : Entrée du *Lexique de cooccurrents — Bourse et conjoncture économique* (Cohen 1986 : 54)

Emprunt : Dette contractée auprès d'agents économiques, sur le marché des capitaux, par l'État, une collectivité publique ou semi-publique ou une entreprise privée à des conditions fixées d'avance.				
	NOMS	VERBES (SUJET)	VERBES (OBJET)	ADJECTIFS
DÉBUT	émission lancement		émettre lancer ouvrir	
CROISSANCE	accroissement augmentation	s'accroître augmenter monter	accroître augmenter	considérable élevé gros
INDÉTERMINÉS				
DÉCLIN	baisse diminution réduction	baisser diminuer	réduire restreindre	petit
FIN			clore liquider rembourser restituer	
AUTRES COOCCURRENTS			contracter obtenir souscrire consolider	

Les travaux de Cohen montrent qu'il est possible de systématiser les relations sémantiques entre termes et cooccurrents et qu'un formalisme comme les FL est utilisable pour en rendre compte.

1.2.2 Laporte (1996, 1997)

Dans le domaine médical, Laporte (1996) a élaboré une méthode, inspirée également des travaux de Mel' uk *et al.* (1984, 1988, 1992, 1995), pour recenser et présenter des combinaisons lexicales observées dans le domaine. Afin d'illustrer cette méthode, l'auteure a effectué le recensement et la consignation des combinaisons lexicales spécialisées concernant une quinzaine de mots clés du domaine de la pharmacologie cardiovasculaire (1997). Laporte a élaboré un outil informatique qui se présente sous forme d'article dans lequel ont été recensés divers renseignements d'ordre collocationnel (les «catégories sémantiques») ainsi que la fréquence et la répartition des combinaisons dans le corpus étudié.

Les combinaisons retenues sont du type terme + verbe (ou verbe + terme), *déclencher une maladie*; terme + adjectif (ou adjectif + terme), *forte dose*; nom + préposition + terme, *défaillance d'une fonction.* Les catégories, au nombre de dix (10), ont été établies en fonction de constantes sémantiques et elles se définissent comme suit, accompagnées d'exemples :

1. faible intensité ou diminution de l'intensité (- intensité) : *posologie prudente*;
2. forte intensité ou augmentation de l'intensité (+ intensité) : *forte dose*;
3. création ou activité (créact) : *engendrer une maladie*;
4. création ou activation inhibée (anticréact) : *inhibition de l'action*;

5. fin ou disparition (fin) : *abandonner un médicament*;

6. évaluation positive (pos) : *myocarde efficace*;

7. évaluation négative (neg) : *sombre pronostic*;

8. courte durée ou durée raccourcie (- durée) : *action fugace*;

9. longue durée ou durée allongée (+ durée) : *maintenir un traitement*.

Une dixième catégorie rendant compte de la neutralité (neu) s'ajoute pour classer des cooccurrents ne présentant pas de particularité sémantique.

Nous reproduisons à la figure 6 ci-dessous une partie de l'entrée *dose* telle qu'elle est traitée dans la recherche de Laporte (Laporte : 1996) :

Figure 6 : Entrée *dose*

DOSE : «Quantité de médicament qui doit être administrée en une seule fois» (Robert : 1994).

ADJECTIF

créact de charge d'attaque	**+ durée** d'entretien (- intensité)
+ intensité forte (hab. prép.) élevée haute (hab. prép.) importante	**- intensité** faible (hab. prép.) modérée d'entretien (+ durée)
pos thérapeutique efficace	

Laporte a démontré que, dans le domaine de la pharmacologie cardiovasculaire, il était également possible de systématiser les relations sémantiques entre les termes et les cooccurrents.

1.2.3 Pavel (1993, 1994)

Pavel a publié un *Guide de la recherche phraséologique en langue de spécialité* (1994) dans lequel elle a élaboré une typologie de ce qu'elle nomme les *phraséologismes*. Son ouvrage se veut essentiellement un guide pour les terminologues dont la vocation consiste à compléter la méthodologie de recherche thématique et ponctuelle en y intégrant une composante phraséologique. Elle définit le phraséologisme comme «une combinaison préférentielle ou une solidarité lexicale d'un terme, appelé noyau ou base, et de mots qui cooccurrent dans la même phrase (cooccurrents, mots tandem)» (Pavel 1994 : 4).

De plus, Pavel distingue : i) la combinatoire syntagmatique, qui relève de la séquence horizontale des cooccurrents dans la phrase (par exemple, l'*arête* + change + de visibilité); et ii) la combinatoire paradigmatique, qui liste des cooccurrents de même catégorie grammaticale pouvant apparaître dans une position déterminée, un à la fois (l'*arête* + change d'état/compose des facettes/relie des sommets) (Pavel 1994 : 19).

Dans sa typologie, l'auteure prend également en compte l'aspect lexico-sémantique et le facteur de commutabilité des éléments d'une unité en regroupant les combinaisons en trois catégories : combinaison lexicalement figée (*promulguer une loi)*, combinaison dont aucun des composants ne peut être remplacé, permuté ou éliminé et à laquelle aucun élément ne peut être ajouté sans que le sens ne change (Pavel 1994 : 7)*;* combinaison à commutabilité restreinte (*déformer par [sous] translation*), combinaison dont au moins un élément lexical peut être remplacé par un synonyme, déplacé ou éliminé, et à laquelle au moins un nouvel élément peut être ajouté sans que le sens global ne change (Pavel 1994 : 7); combinaison libre ([l']*agrégat adopte une configuration, colle à un autre, grossit*, etc.), combinaison d'un terme noyau avec n'importe quel membre d'une famille de mots synonymes en langue générale, prévisible en raison de la compatibilité conceptuelle des composants et dont on peut déduire le sens global en cumulant les sens des parties (Pavel 1994 : 7).

Les travaux de nature pragmatique de Pavel mettent en lumière la complexité liée à l'élaboration d'une typologie structurelle des combinaisons lexicales qui tienne compte de l'organisation lexicale, syntaxique et sémantique des combinaisons.

1.2.4 Martin (1992)

Pour sa part, Martin (1992) décrit les CLS, qu'il appelle *collocations*, comme étant des groupements de mots lexicalement contraints, c'est-à-dire des combinaisons lexicales non libres. À l'opposé des expressions idiomatiques qui sont non compositionnelles, les CLS ont, toujours selon Martin, un sens souvent compositionnel.

Selon l'auteur, la collocation se compose de deux lexèmes dont la relation est dite polaire[14], l'un des constituants (le *collocatif* ou *modificateur*) modifiant l'autre (la *base* ou *tête*).

Par ailleurs, l'étude de Martin se révèle particulièrement intéressante en ceci que l'auteur distingue trois types de structures possibles de combinaisons : 1) lexicales (*célibataire endurci*) (Martin 1992 : 161), dans laquelle l'adjectif *endurci* perd son sens premier en fonction de la base qui exerce ainsi une contrainte d'ordre lexical sur son collocatif; 2) conceptuelles (*maladie infectieuse*) (Martin 1992 : 161), combinaison susceptible de généralisation[15], le terme *maladie* pouvant accepter des cooccurrents sémantiquement apparentés à *infectieuse*; et 3) contextuelles, notion que Martin emprunte à Kjaer (1990), qui a particulièrement étudié des textes juridiques allemands. Selon celle-ci, les relations contextuelles

---

[14] Notons que cette caractéristique est empruntée à Haussmann (1979).
[15] Nous reviendrons sur cette notion de généralisation au point 1.3.

peuvent s'expliquer uniquement dans le cadre du contexte particulier dans lequel elles se trouvent.

1.2.5 Heid (1992, 1994); Heid et Freibott (1991)

Heid a comparé diverses approches de description des collocations (en langue générale) et des combinaisons lexicales spécialisées (en langue de spécialité). En fonction de ses recherches, il s'est ensuite proposé de décrire les collocations de la langue générale en se basant sur trois niveaux descriptifs — conceptuel, lexical et syntaxique (Heid 1993 : 533-534).

Aux niveaux conceptuel ou sémantique, il considère la collocation comme étant une combinaison de deux éléments dont chacun a une sémantique propre et qui n'est pas entièrement analysable sur le plan sémantique par des règles compositionnelles.

Au niveau lexical, Heid décrit la collocation comme étant une combinaison de deux lexèmes, lesquels peuvent être composés de plusieurs mots.

Au niveau syntaxique, il estime, comme Haussmann (1979), que l'on peut distinguer cinq catégories (nom + nom; nom + verbe; nom + adjectif; adjectif + adverbe; verbe + adverbe — se reporter au point 1.1.1). Heid emprunte d'ailleurs à Haussmann (1979) les mêmes caractéristiques linguistiques pour

décrire la collocation (Heid et Freibott 1991 : 79) : i) combinaison «polaire» pour souligner le fait que l'un des deux lexèmes détermine ceux avec lesquels il peut être combiné; ii) distinction entre «base» et «collocatif»; iii) caractère conventionnel de la collocation.

Par ailleurs, Heid et Freibott (1991) se sont intéressés au traitement informatique des combinaisons lexicales dans le cadre de l'élaboration d'un outil informatique en milieu de travail[16]. Ils ont imaginé un outil informatique multifonctionnel, utilisable à la fois comme dictionnaire de thème, dictionnaire de version et dictionnaire unilingue. Cette modélisation de l'information collocationnelle a été conçue dans un but bien précis : fournir aux traducteurs et rédacteurs de textes techniques un outil qui leur donne accès à de l'information de type collocationnel. Cette préoccupation répond à un besoin grandissant parmi les professionnels langagiers, celui de connaître l'environnement phrastique entourant un terme donné de manière à produire des textes ou des traductions fidèles au «registre» avec lequel communiquent les spécialistes scientifiques ou techniques.

---

[16] Notamment pour le service de traduction et de documentation de l'entreprise KRUPP Industrietechnik GmbH, à Duisburg, en Allemagne.

Pour résumer cette section relative au point de vue de la langue spécialisée, retenons les observations suivantes :

1. les travaux de Cohen (1983, 1986) et de Laporte (1996, 1997) démontrent qu'il est possible de systématiser les relations sémantiques entre les termes et les cooccurrents des CLS de certains domaines, ce qui se révèle extrêmement utile pour confectionner des outils d'ordre terminologique dans lesquels on peut rendre compte, en procédant par généralisations, des CLS en LS;

2. il semble que dans tous les cas, la «base» de la CLS soit un terme, généralement de nature nominale;

3. Martin et Heid soulignent également que les CLS, à l'instar des collocations en LG, sont des groupes de mots lexicalement contraints qui entretiennent une relation polaire;

4. les CLS ont parfois un sens semi-compositionnel, mais cette propriété ne caractérise pas l'ensemble des combinaisons en LS.

1.3 Collocation lexicale vs collocation conceptuelle

Comme nous avons pu le constater à la lumière des diverses études des chercheurs en LS, il semble que les CLS se prêtent bien à une analyse fondée sur

des constantes sémantiques. Certains ont poussé plus loin leurs recherches en se basant sur l'hypothèse voulant que, étant donné que les CLS rendent compte d'expressions propres à des domaines de spécialité, on y retrouverait probablement davantage de collocations conceptuelles (dont le cooccurrent peut se combiner avec une classe d'unités terminologiques) que de collocations lexicales (dont le cooccurrent ne s'emploie qu'avec un seul terme). Soulignons toutefois que certains chercheurs en LG, notamment Mel' uk et Wanner (1996) (voir le point 1.3.1.1 ci-après), ont remarqué que le procédé de généralisation conceptuelle n'était pas exclusif à la langue spécialisée.

1.3.1 Langue générale

1.3.1.1 Mel' uk et Wanner (1996)

Mel' uk et Wanner ont mené une étude sur les noms allemands dénotant une émotion (par ex., *Achtung* «respect»; *Angst* «peur»; *Arger* «colère»). Ils ont sélectionné 40 noms qui étaient les mots clés autour desquels s'articulaient dix (10) FL verbales. À partir de la valeur obtenue, 25 des verbes ont par la suite été retenus pour former des combinaisons verbe + nom (par ex., *éprouver de la compassion*). Les auteurs ont décrit les noms au moyen d'un certain nombre de traits (la méthode utilisée est illustrée ci-après).

Figure 7 : Noms dénotant une émotion et leurs composantes sémantiques

Trait sémantique	***Achtung* «respect»**	***Angst* «peur»**	***Arger* «colère»**
Intensité			
Polarité	+	—	—
Manifestation	—	—	—
Direction	+	+	+
Mentalité	+		
Réactivité		+	+
Attitude	+		
Activité	+	+	
Excitation		+	+
Maîtrise de soi		+	
Permanence	+	±	—

Dans leur recherche, les auteurs ont retenu onze composantes sémantiques. Ces composantes comportent de une à trois valeurs mutuellement exclusives, que nous reproduisons dans la figure 8 ci-après :

Figure 8 : Valeurs des composantes sémantiques

1.	Intensité	= intense, modérée, neutre
2.	Polarité	= agréable, désagréable, neutre
3.	Manifestation	= manifestée, manifestable, neutre
4.	Direction	= dirigée vers, neutre
5.	Mentalité	= mentale, neutre
6.	Réactivité	= réaction, neutre
7.	Attitude	= dirigée vers quelqu'un, neutre
8.	Activité	= active, neutre
9.	Excitation	= état d'excitation, neutre
10.	Maîtrise de soi	= perte de la maîtrise de soi, neutre
11.	Permanence	= temporaire, permanente

Pour chacun des mots clés, les auteurs ont précisé la valeur de toutes les composantes, qu'ils ont notée par les marqueurs «plus» (+) (qui désigne la première valeur) ou «moins» (—) (qui désigne la valeur opposée, la deuxième valeur). Ils ont laissé en blanc les composantes de valeur neutre. Ainsi, par exemple, *Achtung* «respect», *Angst* «peur» et *Arger* «colère» partagent tous les trois la composante sémantique «direction» positive, c'est-à-dire dont la valeur indique que l'émotion ressentie est dirigée vers quelqu'un ou quelque chose; d'autre part, les mots clés *Angst* «peur» et *Arger* «colère» partagent la composante «excitation» positive, qui dénote un état d'excitation chez le sujet qui la ressent.

Ils remarquent que, dans certains cas, il est possible de généraliser les fonctions lexicales à des ensembles de mots clés en fonction des composantes sémantiques de ces derniers.

> *The treatment of lexical data [...] shows that significant correlations between restricted lexical co-occurrence and semantic features exist, and they allow for reasonable generalizations. At the same time, the correlations are far from absolute: idiosyncracies in collocations abound and simply have to be listed.* (Mel' uk et Wanner 1996 : 211)

Cependant, comme on le constate, les généralisations ne peuvent être pratiquées pour l'ensemble des FL.

À notre connaissance, il n'existe pas d'autres études de ce genre qui portent sur les collocations de la langue générale.

1.3.2 Langue spécialisée

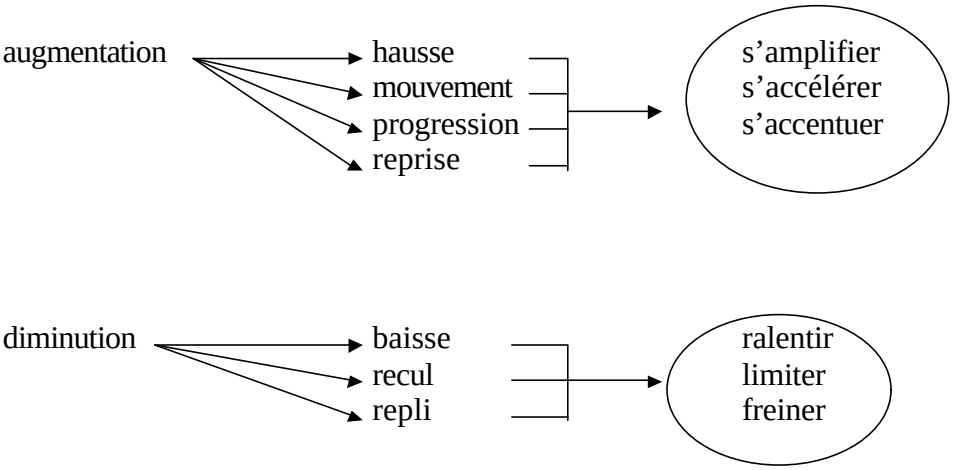
1.3.2.1 Heid (1992, 1994)

Pour ce qui est des combinaisons en langue spécialisée, Heid (1992, 1994) avance que le niveau conceptuel revêt une importance particulière et qu'il est possible de prévoir des combinaisons lorsque l'on se trouve en présence d'une description conceptuelle détaillée. Il déclare également qu'il y a plus de chances de prévoir les combinaisons possibles en LS étant donné que les termes de

beaucoup de domaines sont moins polysémiques que les mots de la langue générale.

Heid a étudié, entre autres, le *Lexique de cooccurrents — Bourse et conjoncture économique* (Cohen : 1986), dans lequel il a relevé les noms partageant les mêmes cooccurrents. Il a remarqué, par exemple, que les noms dénotant une «augmentation» ou une «diminution» — *hausse*, *baisse*, *mouvement*, *progression*, *recul*, *repli*, *reprise* — partageaient, d'une part, pour le concept d'«augmentation» les verbes : *s'amplifier*, *s'accélérer*, *s'accentuer*, et, d'autre part, pour le concept de «diminution» les verbes : *ralentir*, *limiter*, *freiner*. Le schéma ci-dessous illustre les observations de Heid :

Figure 9 : Regroupement des noms dénotant une augmentation ou une diminution



Cette constatation a conduit Heid à mettre en relation le niveau de description conceptuel et le comportement collocationnel, soulignant le caractère conceptuel de certains ensembles et sous-ensembles de CLS.

Notons toutefois que cette étude, basée sur un ouvrage existant et non sur un corpus composé de textes, ne peut prétendre couvrir de manière exhaustive l'ensemble du domaine économique. Elle signale la présence de collocations conceptuelles, mais n'en mesure pas l'importance.

1.3.2.2 Heid et Freibott (1991)

Heid et Freibott, à l'instar de Martin (1992) (voir le point 1.2.4), ont soulevé plusieurs questions pertinentes quant à la nécessité de distinguer les collocations lexicales et les collocations conceptuelles. Ils suggèrent également qu'il existe une prédominance des combinaisons conceptuelles dans les domaines spécialisés, en raison de la spécificité liée à la dénomination de notions concrètes, par opposition à une prédominance des collocations lexicales dans la langue générale. Encore une fois, ces observations n'ont pas fait l'objet de vérifications systématiques.

1.3.2.3 Meyer et Mackintosh (1994, 1996)

Pour leur part, Meyer et Mackintosh signalent que certains domaines se prêtent particulièrement bien à une analyse terminologique fondée sur une approche conceptuelle. En effet, Meyer et Mackintosh ont remarqué que certaines CLS, particulièrement dans le domaine des disques optiques, peuvent être généralisées à des ensembles de mots clés (ou bases) appartenant à la même classe conceptuelle.

Meyer et Mackintosh, comme les auteurs qui ont fait l'objet des sections précédentes, constatent la présence de collocations conceptuelles sans pousser plus loin l'examen de celles-ci.

1.3.2.4 L'Homme (1993, 1995, 1998); L'Homme et Gemme (1997)

Les recherches de L'Homme (1993, 1995, 1998) portent particulièrement sur le recensement et la consignation des CLS dans des outils informatisés. Pour ce faire, l'auteure a étudié et retenu les propriétés et caractéristiques formulées par d'autres spécialistes, dont plusieurs ont fait l'objet de la présente étude, à savoir le caractère conventionnel des groupements, leur forme, leur nature compositionnelle ou non compositionnelle, la possibilité de généraliser les bases ou mots clés à des ensembles plus vastes d'unités lexicales et les possibilités de généraliser des relations sémantiques entre les éléments des groupements. Ses travaux ont porté

notamment sur l'encodage des combinaisons lexicales en langue de spécialité basé sur une hiérarchie de concepts génériques, l'accent étant ainsi mis sur les relations de type conceptuel de ces combinaisons (de type verbe + terme; déverbal + préposition + terme).

L'Homme s'est d'abord penchée sur les combinaisons verbales de type verbe + terme (par exemple, *release screw*). Elle a imaginé de mettre en place une interface entre une base de données terminologique existante et une base de données verbale. Pour chaque terme faisant l'objet d'une fiche dans une base de données terminologique donnée, on associe une étiquette renvoyant à une classe conceptuelle. Dans la base de données verbale associée, on spécifierait pour chaque verbe la classe conceptuelle de ses compléments et de son sujet, c'est-à-dire le ou les concepts génériques auxquels ils peuvent être associés. Ainsi, dans une fiche portant sur le terme *hexagonal screw*, on spécifierait que cette notion appartient à la classe conceptuelle des *éléments de fixation*. Puis, dans la fiche portant sur le verbe *release*, on spécifierait que celui-ci peut avoir comme complément d'objet des noms appartenant à la catégorie sémantique des *éléments de fixation*.

Par exemple, supposons l'interrogation suivante portant sur la combinaison lexicale *absorb used waters* (L'Homme 1995 : 157) :

Figure 10 : Combinaison ABSORB + USED WATERS

ABSORB + USED WATERS
nothing
used waters < water

ABSORB + WATER
nothing
water < liquid

1. ABSORB (liquid, gas)
   Def. To draw.
   absorber (un liquide, un gaz)

Le système rechercherait d'abord la chaîne de mots *absorb + used waters*. La recherche étant infructueuse, il remonterait alors à un concept générique, celui de *water* et rechercherait la chaîne *absorb + water.* Si on n'aboutit à aucun résultat, le système remonterait encore vers un concept générique, *liquid* cette fois, et il trouverait finalement une occurrence de *absorb + liquid*, avec la combinaison équivalente en français.

Dans le domaine informatique, L'Homme et Gemme (1997) ont procédé de manière semblable. Au moyen d'un codage des classes conceptuelles auxquelles appartiennent les termes, les auteurs ont établi un modèle de base de données comprenant des enregistrements terminologiques et des enregistrements verbaux. Nous présentons un exemple de ce modèle ci-dessous (L'Homme et Gemme 1997 : 92) :

Figure 11 : Entrée d'une base de données terminologique incluant un enregistrement verbal

ENREGISTREMENT TERMINOLOGIQUE		ENREGISTREMENT VERBAL	
Entry	data file	Entry	delete
Definition	A file containing data records	Structure	V + O
Entrée	fichier de données	Definition	To remove something that has previously been written or stored.
Définition	Fichier qui comporte un ensemble de données	Subject	PEOPLE
Domaine	informatique	Object	ENT REP
Concept	ENT REP	Entrée	effacer
		Structure	V + O
		Définition	...
		Sujet	PEOPLE
		Objet	ENT REP

Dans cet exemple, l'entrée *data file* entre dans une classe conceptuelle regroupant des objets «destinés à diffuser ou à contenir de l'information». Ces unités sont étiquetées comme des *representational entities* (ENT REP). Quant au verbe *effacer*, on indique qu'il peut admettre des sujets appartenant à la classe des animés humains (PEOPLE, par exemple, *user*, *programmer*). Le verbe admet également des objets appartenant à la classe des *representational entities*. Ce modèle permet ainsi de repérer la combinaison «effacer fichier de données» ou bien la combinaison «programmeur efface (fichier de données)».

Les divers travaux de L'Homme lui ont permis de dégager certaines constantes relativement à la nature des collocations et des combinaisons lexicales spécialisées. L'auteure signale les observations suivantes :

1. le principe de généralisation semble productif en langue de spécialité;
2. les groupements compositionnels semblent plus fréquents en langue de spécialité et ils peuvent être généralisés.

Notons que les travaux de L'Homme mentionnés dans le présent mémoire, qui tiennent pour acquis que les CLS sont généralisables, n'ont pas vérifié à partir de dépouillements systématiques de corpus dans quelle mesure les collocations conceptuelles étaient importantes. Le système d'interrogation qu'elle a élaboré est conçu pour décrire des collocations qu'elle considère *a priori* comme conceptuelles.

De la même façon, Heid se fonde sur différents travaux, notamment une étude menée sur l'ouvrage de Cohen (1986), pour constater l'existence de collocations lexicales et de collocations conceptuelles sans toutefois en mesurer l'importance.

En somme, plusieurs auteurs ont constaté l'existence de collocations lexicales et de collocations conceptuelles en langue de spécialité, mais sans analyser de manière approfondie leur comportement ni leur importance.

Notre démarche diffère de celle des auteurs précités en ceci que nous baserons nos observations sur un corpus regroupant plusieurs textes des domaines concernés. Notre recherche portera sur un ensemble de combinaisons lexicales

que nous aurons recensées dans notre corpus, ce dernier reflétant les habitudes langagières de spécialistes.

Nous présentons au chapitre deux la méthodologie que nous avons adoptée pour mener à bien cette recherche.

CHAPITRE II : MÉTHODOLOGIE

Ce chapitre présente les différentes étapes de notre recherche, laquelle se fonde sur un corpus que nous avons constitué à partir d'ouvrages sélectionnés dans les deux domaines à l'étude (voir les annexes 1 et 2). Notre propos étant de faire ressortir les différences de comportement des combinaisons lexicales dans deux discours spécialisés, il nous a fallu élaborer une base de données à deux volets, et, partant, gérer les données recueillies en parallèle.

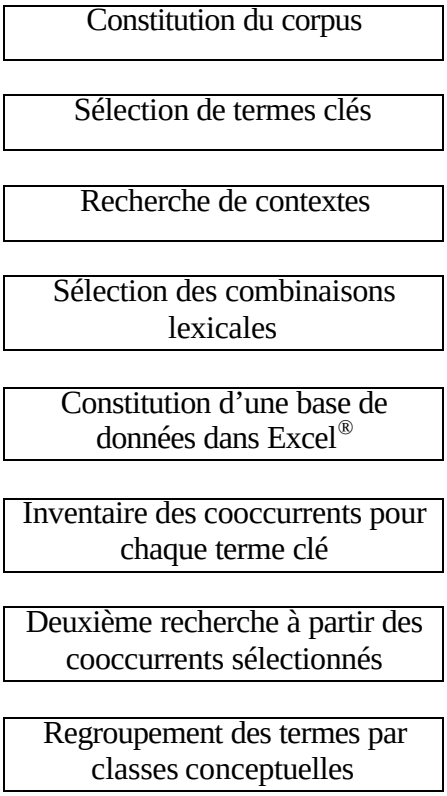
Pour faciliter notre recherche, nous avons opté pour l'utilisation de divers outils informatiques tels que le lecteur optique et, outre le logiciel de traitement de textes Word®, le logiciel tableur Excel® ainsi que le logiciel de gestion documentaire Naturel Pro^MC^. Nous présenterons ces logiciels au moment opportun.

À partir du corpus établi, nous avons d'abord sélectionné des termes clés représentatifs de chacun des domaines concernés, et nous avons ensuite répertorié les nombreuses combinaisons lexicales en fonction des éléments qui les composaient. Puis, à la suite d'un second inventaire de cooccurrents, nous avons tenté de regrouper les CLS obtenues par classes conceptuelles, le cas échéant, afin d'observer les collocations lexicales par rapport aux collocations conceptuelles.

2 Processus de recherche

Le diagramme de cheminement ci-après illustre les différentes étapes que nous avons suivies et que nous nous proposons d'expliquer en détail aux points suivants.

Figure 12 : Diagramme de cheminement



2.1 Constitution du corpus

2.1.1 Choix des domaines étudiés

En tant que traductrice, nous avons été appelée à traduire des textes extrêmement variés dont les discours comportaient d'intéressants défis, entre autres d'ordre terminologique. Nous nous sommes donc basée sur notre expérience personnelle pour fixer notre choix sur deux domaines de spécialité qui nous semblaient suffisamment distincts pour pouvoir en observer les différences de comportement lexical. Ainsi, l'un, l'aéronautique, est associé aux sciences dites exactes, et l'autre, la philosophie, est associé aux sciences humaines.

2.1.2 Ouvrages portant sur le domaine de l'aéronautique

Science de la navigation aérienne, l'aéronautique est un domaine fort vaste. Nous avons sélectionné des textes liés au contrôle aérien, qui englobe les notions spécifiques aux aéroports, au pilotage, et, bien sûr, à l'avion, entité centrale de cette sphère de l'activité humaine. Les ouvrages, au nombre de sept, reflètent, croyons-nous, le caractère technique du vocabulaire utilisé par les «gens de l'air». On trouvera à l'annexe 1 une liste des ouvrages retenus. En outre, pour les besoins de notre étude, nous avons inclus dans notre corpus (en recourant au procédé de numérisation, voir le point 2.1.6 ci-après) uniquement certains chapitres extraits en fonction de leur pertinence.

2.1.3 Ouvrages portant sur le domaine de la philosophie

Autre domaine des plus étendus, la philosophie se définit comme suit :

> «Ensemble des études, des recherches visant à saisir les causes premières, la réalité absolue ainsi que les fondements des valeurs humaines et envisageant les problèmes à leur plus haut degré de généralité.» (Robert 1995 : 1660)

Parmi les grands thèmes de la philosophie, on peut citer, entre autres, l'esthétique, la logique, la métaphysique, la morale. Nous avons choisi d'étudier plus particulièrement la métaphysique, notamment les notions de l'être, de l'âme et de la spiritualité. Encore une fois, on trouvera à l'annexe 2 une liste des ouvrages retenus dans ce domaine.

2.1.4 Choix de la langue

Notre analyse vise uniquement la combinatoire française. Par conséquent, les documents retenus sont ceux rédigés préférablement en français, d'origines diverses mais de même nature, par des spécialistes, pour un public de spécialistes du même domaine. Ainsi, pour ce qui est du domaine de l'aéronautique, nous avons conservé principalement les ouvrages de nature normative et éducative. Dans le domaine de la philosophie, les ouvrages sont d'ordre plutôt général et d'un niveau de spécialisation intermédiaire, présentant la pensée de quelques philosophes dans ses grandes lignes.

2.1.5 Échantillonnage

La taille globale du corpus s'élève à 282 849 mots. La répartition des données en fonction de chaque domaine étudié se détaille de la façon suivante :

- aéronautique : 159 373 mots;
- philosophie : 122 476 mots.

Nous avons jugé que la taille du corpus était suffisante. En effet, comme on le verra plus loin, nous avons recueilli un nombre assez élevé de CLS, dont chacune est attestée par une fréquence minimale.

2.1.6 Numérisation

2.1.6.1 Lecture optique des documents

Nous avons choisi de procéder à une numérisation des documents de référence pour permettre leur traitement ultérieur par un logiciel de recherche de contextes (Naturel Pro^{MC}).

La numérisation est un «processus effectué par les lecteurs optiques et consistant en la conversion des images en données numériques (suites de bits) en vue de leur conservation sur support numérique ou optique et de leur traitement par ordinateur» (OLF : 1996).

L'étape de numérisation, somme toute mécanique, a exigé beaucoup de temps. Il suffit de mentionner, par exemple, que la numérisation d'un document d'une centaine de pages peut prendre environ trois (3) heures. L'opération consiste essentiellement à «balayer» le texte à l'aide d'un lecteur optique, appareil qui fonctionne un peu à la manière d'un copieur. Le résultat peut ensuite être manipulé par l'utilisateur. En effet, à l'aide d'un logiciel de reconnaissance optique de caractères, il est possible de sauvegarder le texte dans n'importe quelle application de traitement de texte (WordPerfect, Word, Write, etc.), et de l'éditer à sa convenance par la suite.

Cependant, toute sophistiquée qu'elle soit, la technologie moderne est loin d'être parfaite. En effet, la qualité du document numérisé est directement proportionnelle à celle du document original. Ainsi, dans le cas de manuels récents, dont la police de caractères est classique et dont le corps se situe entre 10 et 12 points, ou dont les pages ne sont abîmées d'aucune manière, le résultat est acceptable. Par contre, dans le cas de manuels plus anciens, dont les pages sont jaunies ou froissées, dont la police de caractères est petite ou d'un style particulier, ou dans le cas de revues dont les pages sont glacées, par exemple, le résultat est loin d'être satisfaisant. Il est alors essentiel de procéder à une correction orthographique à l'aide d'un logiciel de traitement de texte si l'on désire utiliser le document à des fins de recherche informatisée.

2.1.6.2 Correction des documents

À la figure 13 ci-dessous, nous présentons un exemple de document numérisé. Nous avons utilisé le soulignement double pour mettre en évidence les erreurs de reconnaissance de caractères :

Figure 13 : Exemple de texte numérisé

Celsius Systeme de compte rendu quohdien des événements de l'Aviahon civile Extension de region de conhôle CanadaAirPilot Station radio d'aerodrome communautaire Turbulence en ciel clair Catégone 1,11,111 Plafond et visibilité OK Cumulonimbus Procédures d'inspection d'état et de confonmité Conseil consultabf sur la réglementahon aérienne canadienne Espace aénen inténeur canadien Cerbficatd~immahiculabon Cerbficat de navigabilite Services de secours et d'incendies d'aéropon (MDN) Supplément de vol - Canada Centre de gravité Cenbe météorologique du Canada Specificabons canadiennes dc perfonmanees minimales de navigabon Espaoe aénen canadien à specificabons de perfonmances minimales de navigation Point de transibon Codificabon des reglements du Canada Région de contrôle VFR contrôlé Zone de contrôle

Le document original est récent (1995)[17], mais la taille des caractères est en deçà de 10 points. On remarque que le logiciel ne lit pas correctement certains caractères, notamment les «tr» et les «rt» qu'il rend en «h»; les «ti» en «b»; les «ri» en «n»; les «e» en «c», les accents, etc. Le traitement informatisé de tels

[17] Extrait de «Publication d'information aéronautique», A.I.P. Canada.

documents se révèle alors difficile, et la recherche de termes presque impossible : un outil informatisé de recherche terminologique, aussi puissant soit-il, ne pourra pas retrouver la chaîne de caractères correspondant au syntagme «espace aérien» dans le cas présent. L'«oeil» humain saura reconnaître et substituer les bonnes lettres à la chaîne «espace aénen», mais pas la machine.

Ainsi, avant d'effectuer tout traitement informatisé, il est impératif de corriger les erreurs que le logiciel de reconnaissance de caractères a glissées dans le document. Pour ce faire, on a recours à un traitement de texte, qui, grâce à la fonction de correction orthographique, allège de beaucoup cette tâche des plus fastidieuses, voire titanesque au regard d'un corpus de plus de 200 000 mots.

Une fois cette étape terminée, nous avons procédé à la recherche de contextes à l'aide du logiciel Naturel Pro^{MC} pour Windows décrit ci-après au point 2.3.1.

2.2 Sélection des termes clés

Pour chacun des domaines, nous avons choisi cinq (5) termes clés qui ne sont pas sémantiquement apparentés (auxquels nous avons associé une étiquette opératoire uniquement aux fins de la présente recherche et dont il sera question au point 2.8.2), et qui dénotent des notions importantes dans les domaines concernés.

En outre, les termes ne devaient pas être sémantiquement apparentés afin de varier les cooccurrents et de vérifier s'il était possible de créer un certain nombre de classes conceptuelles.

Le tableau I ci-après présente la liste des termes clés sélectionnés. Soulignons que notre recherche s'est faite d'abord à partir des termes, à l'instar de la méthode utilisée par les terminologues.

Tableau I : Termes clés sélectionnés

DOMAINE	TERME CLÉ	ÉTIQUETTE OPÉRATOIRE
Aéronautique	• aéronef • aéroport • piste • vitesse • vol	➢ appareil volant ➢ installations propres au trafic aérien ➢ voie de circulation ➢ unité de mesure ➢ activité
Philosophie	• amour • beauté • connaissance • être • vérité	➢ disposition ➢ propriété ➢ entité morale ➢ entité abstraite ➢ principe

2.3 Recherche de contextes

Dans le but de sélectionner des combinaisons lexicales représentatives de chacun des domaines, nous avons procédé à une recherche préalable de contextes. Pour ce faire, nous avons eu recours au logiciel Naturel Pro^{MC}, que nous présentons au point suivant.

2.3.1 Présentation du logiciel Naturel Pro^{MC}

Naturel Pro^{MC} pour Windows est, en fait, un logiciel de gestion et de recherche d'information textuelle. La base documentaire peut être constituée de documents de formats différents, créés par traitement de texte, base de données ou chiffrier électronique. Nous avons choisi de travailler avec ce logiciel car, grâce à son mode de recherche plein texte et à sa capacité d'indexation positionnelle, il nous a permis d'interroger aisément un ensemble de fichiers et de repérer sans peine des contextes dans le but d'illustrer les combinaisons lexicales ou les termes choisis.

On crée d'abord une base en constituant une liste des fichiers voulus. Puis, Naturel indexe le contenu de ces fichiers, que l'on peut alors interroger soit en langage naturel, soit à l'aide d'opérateurs booléens. En outre, Naturel possède une interface à affichage dynamique constituée de quatre (4) fenêtres présentant les résultats de la recherche sous diverses formes (figure 14 ci-après) :

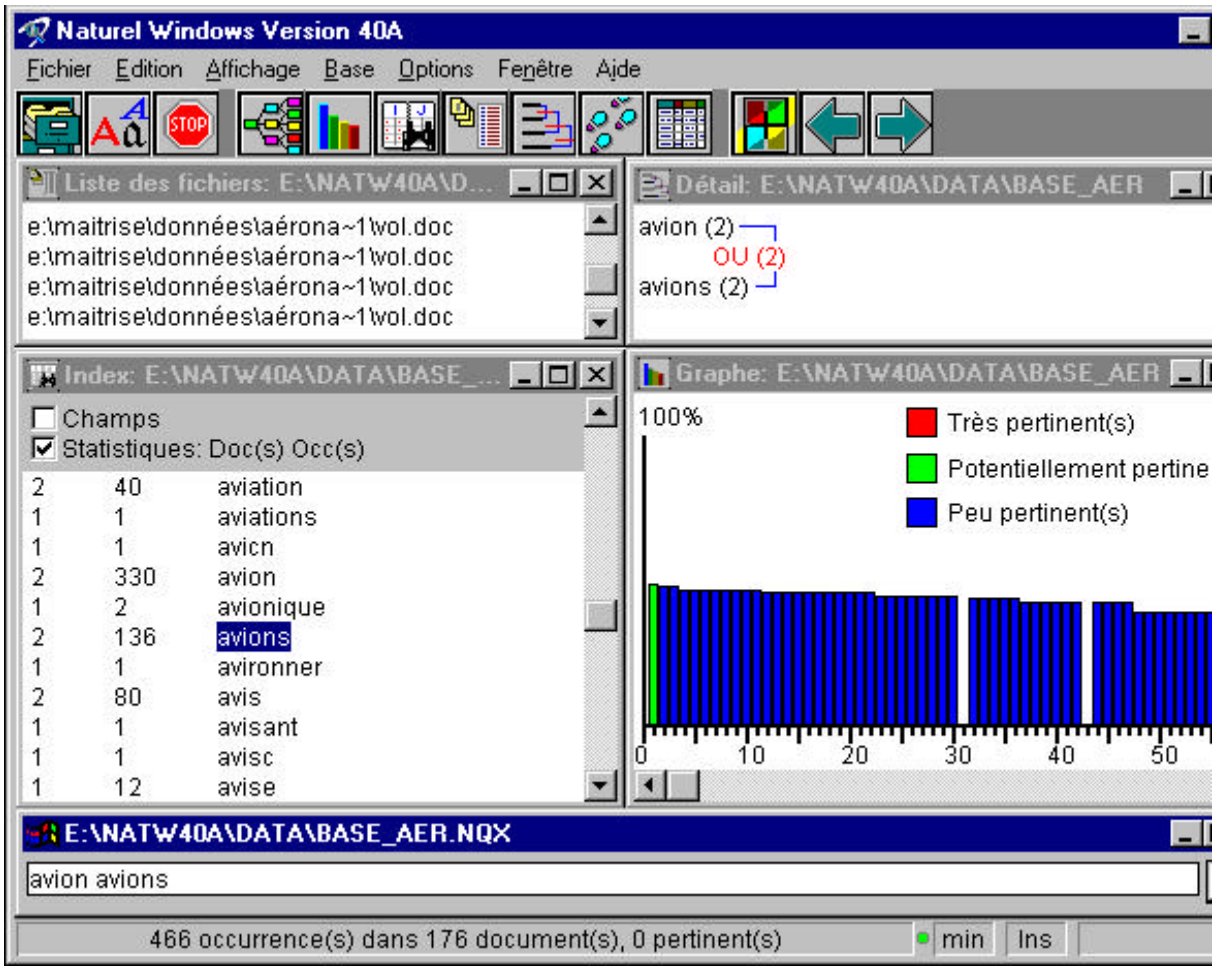
1. une liste des documents qui répondent à la question posée par ordre de pertinence, c'est-à-dire, en ce qui concerne la recherche de contextes, selon le nombre d'occurrences dans un segment donné;

2. un graphique à barres de ces documents par ordre de pertinence également (Naturel classe les documents selon leur pertinence : très pertinent, potentiellement pertinent et peu pertinent);

3. un index des mots disponibles dans la base;

4. le détail de la recherche sous forme hiérarchique.

Figure 14 : Écran de saisie des quatre fenêtres d'affichage dynamique proposées par Naturel Pro[MC]



Cette saisie d'écran nous montre le résultat de la recherche portant sur le terme *avion*. La zone d'interrogation est située au bas de l'écran. Dans le cas présent, nous avons inclus dans notre question les formes singulière et plurielle du terme recherché. Le logiciel nous présente alors le résultat : 466 occurrences dans 176 documents (ou segments). Il suffit de cliquer ensuite soit sur l'un des segments reproduits sur le graphique, soit directement dans la liste des segments

présentée dans la fenêtre Liste des fichiers. On accède alors au segment de fichier contenant le terme recherché. Ce dernier est mis en surbrillance pour permettre une visualisation plus aisée. Il est alors possible de copier et coller les contextes retenus soit dans une application de traitement de texte, soit directement dans une base de données.

2.4 Sélection des combinaisons lexicales spécialisées (CLS)

2.4.1 Définition

Rappelons brièvement la définition que nous avons adoptée : par «combinaison lexicale spécialisée» nous entendons toute combinaison de deux lexèmes dont la base (ou noyau) constitue une unité terminologique (simple ou complexe – soit un nom). Rappelons que les bases retenues sont celles présentées au tableau I.

Les unités lexicales qui cooccurrent avec le terme peuvent appartenir aux catégories nom — *rapport de pilote,* verbe — *exploiter [une] ligne aérienne* ou adjectif — *paix intérieure*. De plus, les CLS respectent des conventions d'usage propres à la langue de spécialité concernée.

2.4.2 Terme complexe et CLS

La délimitation des termes complexes n'est guère aisée, et la distinction entre terme complexe et CLS se révèle particulièrement difficile en présence d'associations de types «terme + préposition + nom», «nom + préposition + terme» ou «nom + adjectif».

En terminologie, le terme ou l'unité terminologique complexe se distingue par le fait de désigner une notion de façon univoque, alors qu'une CLS, en langue de spécialité, peut faire appel à un nombre *x* de notions. Par exemple, *joint d'étanchéité*[18] constitue un terme en ce qu'il désigne une seule notion.

> «...la suite d'éléments linguistiques *joint d'étanchéité* est un terme [...] puisqu'elle est une unité signifiante constituée de plusieurs mots qui désigne une notion de façon univoque à l'intérieur du domaine de l'automobile» (Blais : 1993).

En revanche, nous considérons que *espace aérien inférieur* n'est pas une unité terminologique complexe (le principe d'univocité n'étant plus valide), mais plutôt une CLS; une combinaison lexicale privilégiée dans un domaine particulier, soit l'aéronautique. En effet, *espace aérien inférieur* est une combinaison d'éléments linguistiques qui ne désigne pas une notion de façon univoque, mais

---

[18] Pièce de forme, d'épaisseur et de matériau appropriés interposée entre deux surfaces serrées l'une contre l'autre, afin d'assurer l'étanchéité statique d'un assemblage. (OLF : 1996)

une combinaison des notions de «espace aérien», lequel constitue le terme noyau de la combinaison, et de «inférieur».

Il existe cependant des critères de lexicalisation (existence d'une définition spécialisée; position synonymique; maniabilité syntagmatique; cohésion syntaxique, pour n'en nommer que quelques-uns) qui peuvent faciliter le découpage des termes complexes. Ainsi, Kocourek (1982) caractérise le terme complexe comme étant une combinaison d'éléments en une unité plus élevée et appartenant à un système paradigmatique des unités lexicales.

De plus, toujours selon Kocourek, le terme complexe est «unifié par son sens et par ses propriétés dénominatives (dénomination, sémantique, signification)» (1982).

Par ailleurs, selon Gross (1988), la relation qui existe entre les éléments composant un terme complexe présente également un certain degré de figement qu'il est possible de calculer en étudiant les propriétés syntaxiques caractéristiques de cette relation.

L'application de ces critères ne se fait pas toujours sans problème. Pour éviter les situations de conflit entre les CLS et les termes complexes, notamment pour ce qui est des constructions du type nom + préposition + nom ou nom + adjectif, nous avons écarté les unités terminologiques complexes, lesquelles

apparaissent en encadré à l'annexe 8 et ne sont pas comptabilisées dans l'analyse de nos résultats. Pour ce faire, nous avons consulté différents glossaires et vocabulaires de terminologie relatifs aux domaines étudiés en plus de nous fonder sur notre propre connaissance des sujets traités.

2.4.3 Formes de CLS retenues

Nous avons classé les formes de CLS retenues en trois grandes catégories, lesquelles se subdivisent en sous-catégories en fonction du rôle joué par la base de la combinaison.

Tableau II : Catégories de combinaisons lexicales spécialisées

CATÉGORIES	**AÉRONAUTIQUE**	**PHILOSOPHIE**
Combinaison adjectivale :		
Terme + adjectif	*aéroport* contrôlé	*être* empirique
Adjectif + terme	petit *aéroport*	immuable *vérité*
Combinaison nominale :		
Terme + préposition + nom	*trajectoire* de départ	*être* de l'âme
Nom + préposition + terme	circulation d'*aéroport*	quête de l'*être*
Combinaison verbale :*		
Terme + verbe	*aéronef* décolle	*être* répond
Verbe + terme	maîtriser [un] *aéronef*	justifier l'*être*

* (y compris les combinaisons du type «verbe + préposition + terme»)

2.5 Constitution d'une base de données à l'aide du logiciel Excel®

Pour permettre une gestion efficace de l'ensemble des données recueillies, nous avons ensuite constitué une base de données informatisée à l'aide du logiciel Excel®.

2.5.1 Présentation du logiciel Excel®

Excel® se classe parmi les logiciels appelés tableurs. Les tableurs se présentent sous la forme de feuilles composées de lignes et de colonnes et ils permettent de créer des tableaux de différentes natures, notamment des tableaux de calcul de données numériques, d'applications financières, d'études de marché, de graphiques, etc.

Excel® permet également de créer et de gérer des bases de données, un peu à la façon des logiciels spécialement conçus à cette fin tels que Access®, sans toutefois prétendre à la puissance d'organisation et de traitement de ces derniers. En effet, à la différence de Access®, par exemple, Excel® n'est pas une base de données relationnelle (permettant de créer des liens entre des entrées multiples et d'effectuer des requêtes en combinant plusieurs éléments de recherche), et il permet d'exploiter un volume moins important de données. Toutefois, la taille de notre corpus étant plutôt restreinte, nous avons choisi de travailler avec Excel® en raison de certaines de ses fonctionnalités que nous avons jugées des plus pratiques

dans le cadre de notre étude, entre autres la fonction de filtres. En effet, au moyen de ces filtres, nous avons pu retracer aisément et rapidement tous les cooccurrents qui se combinaient avec un terme donné dans les combinaisons particulières recherchées (V + T, T + V, etc.).

2.5.2 Constitution de la base de données

Notre base de données comporte deux volets nécessaires à la présente étude comparative : celui destiné à gérer les données relatives au domaine de l'aéronautique, et celui destiné aux données du domaine de la philosophie. Ainsi, pour chacun des domaines, la base se compose essentiellement de quatre feuilles : la première rassemble toutes les données pertinentes pour l'ensemble des termes clés, soit le terme clé de départ, les cooccurrents composant les diverses combinaisons retenues dans les catégories adjectivales, nominales de même que verbales et les contextes. Nous reproduisons ci-dessous un enregistrement tiré de la base Aéronautique :

Figure 15 : Exemple d'enregistrement de la base Aéronautique[19]

N°	TCD	A	TCA	ATC	N	TCPN	NPTC	V	TCV	VTC	CTC
1	aéroport				contrôleur		O				Le contrôleur d'aéroport peut [...]

[19] Quelques pages de la base de données sont présentées à l'annexe 9.

Dans cet enregistrement, nous distinguons les éléments suivants :

N° : le numéro séquentiel de l'entrée;

TCD : le terme clé de départ;

A : le cooccurrent adjectival du terme clé;

TCA : la place qu'occupe le cooccurrent adjectival dans la combinaison, soit terme clé + adjectif;

ATC : la place qu'occupe le cooccurrent adjectival dans la combinaison, soit adjectif + terme clé;

N : le cooccurrent nominal du terme clé;

TCPN : la place qu'occupe le cooccurrent nominal dans la combinaison, soit terme clé + préposition + nom;

NPTC : la place qu'occupe le cooccurrent nominal dans la combinaison, soit nom + préposition + terme clé;

V : le cooccurrent verbal du terme clé;

TCV : la place qu'occupe le cooccurrent verbal dans la combinaison, soit terme clé + verbe;

VTC : la place qu'occupe le cooccurrent verbal dans la combinaison, soit verbe + terme clé;

CTC : un contexte dans lequel se trouvent le terme clé et son cooccurrent.

La deuxième feuille rassemble les données relatives aux combinaisons verbales relevées à partir du terme clé de départ, soit le verbe cooccurrent, le

terme, la place que le terme occupe dans la combinaison (sujet ou objet) et le contexte (CTX) dans lequel se trouve la combinaison :

Figure 16 : Exemple d'enregistrement verbal[20]

N°	VERBE	TERME	SUJ	OBJ	CTX
7	circuler	aéronef	O		Si plusieurs aéronefs sont en attente, ils circulent sur [...]

La troisième feuille rassemble les données relatives aux combinaisons nominales relevées à partir du terme clé de départ, soit le terme, le nom qui se combine avec celui-ci (COOCCURRENT), la place qu'occupe ce dernier dans la combinaison nominale (antéposition ANTÉ ou postposition POST) et le contexte (CTX) dans lequel se trouve la combinaison.

Figure 17 : Exemple d'enregistrement nominal[21]

N°	TERME	COOCCURRENT	ANTÉ	POST	CTX
119	soi	discernement	O		Présence de l'âme à elle-même: discernement de soi.

La quatrième feuille rassemble les données relatives aux combinaisons adjectivales relevées à partir du terme clé de départ, soit l'adjectif, le terme qui se combine avec celui-ci, la place qu'occupe ce dernier dans la combinaison

[20] Une page de la base de données verbale est donnée en exemple à l'annexe 10.
[21] Une page de la base de données nominale est donnée en exemple à l'annexe 11.

adjectivale (antéposition ou postposition) et le contexte (CTX) dans lequel se trouve la combinaison.

Figure 18 : Exemple d'enregistrement adjectival[22]

N°	ADJECTIF	TERME	ANTÉ	POST	CTX
204	initial	trajectoire	O		Comme tous les autres corps, l'avion est soumis à la loi de l'inertie, c'est-à-dire qu'il tend à continuer de suivre sa trajectoire initiale.

La figure 19 présente un exemple de listes obtenues dans l'un et l'autre domaines. Nous invitons le lecteur à se reporter aux annexes 3 et 4 pour consulter les listes de cooccurrents verbaux, nominaux et adjectivaux que nous avons relevés au cours de notre recherche.

Figure 19 : Listes partielles de cooccurrents

Base Aéronautique — verbes :

atteindre	croiser	franchir	poser
atterrir	décoller	fréquenter	poser (se)
augmenter	dégager	implanter	prescrire
autoriser	déposer	intégrer	régir
calculer	desservir	maintenir	reprendre
circuler	emprunter	maîtriser	survoler
contrôler	exploiter	planifier	traverser

[22] Une page de la base de données adjectivale est donnée en exemple à l'annexe 12.

Base Philosophie — verbes :			
accompagner	conduire	dépasser	émaner
achever	connaître	déprendre (se)	entrer
achever (s')	consentir	détourner (se)	envisager
affirmer	considérer	détruire (se)	évoquer
ajouter	constituer	devancer	faire
apparaître	construire	devenir	fonder
appréhender	correspondre	dire	former

2.5.3 CLS — domaine de l'aéronautique

Sur un total de 3884 CLS, 965 ont été relevées à partir du terme de départ *aéronef*; 565 à partir de *aéroport*; 306 à partir de *piste*; 1 107 à partir de *vitesse*; et 941 à partir de *vol*. Nous avons ensuite comptabilisé les combinaisons verbales, nominales et adjectivales, lesquelles avaient été réparties à leur tour en catégories. Ainsi, pour le terme clé de départ *aéronef*, nous avons constaté que 16 % des CLS se classaient dans la catégorie V + T; 7 % dans la catégorie T + V; 30 % dans celle des N + T; 8 % dans celle des T + N; 0 % dans celle des A + T; et 39 % dans celle des T + A. Comme on peut le constater, les combinaisons nominales et adjectivales sont les plus productives, alors que les combinaisons verbales suivent de près au troisième rang (voir l'annexe 6).

Tableau III : Résultats des combinaisons obtenues — aéronautique

RÉSULTATS_AÉRONAUTIQUE					
TERME	**aéronef**	**aéroport**	**piste**	**vitesse**	**vol**
V + T	158	76	67	80	432
T + V	71	0	0	2	25
TOTAL_VERBE	229	76	67	82	457
N + T	286	81	195	0	349
T + N	77	102	0	804	3
TOTAL_NOM	363	183	195	804	352
A + T	0	115	0	38	0
T + A	373	191	44	183	132
TOTAL_ADJECTIF	373	306	44	221	132
TOTAL_TERME	965	565	306	1107	941

2.5.4 CLS — domaine de la philosophie

Sur un total de 2 301 CLS, 575 ont été relevées à partir du terme de départ *amour*; 255 à partir de *beauté*; 535 à partir de *connaissance*; 496 à partir de *être*; et 440 à partir de *vérité*. Pour le terme clé de départ *amour*, nous avons constaté que 10 % des CLS se classaient dans la catégorie V + T; 7 % dans la catégorie T + V; 9 % dans celle des N + T; 7 % dans celle des T + N; 0 % dans celle des A + T; et 67 % dans celle des T + A. Comme on peut le constater, les combinaisons adjectivales sont les plus productives, alors que les combinaisons verbales et nominales comportent un pourcentage presque égal de combinaisons (voir l'annexe 7).

Tableau IV : Résultats des combinaisons obtenues — philosophie

RÉSULTATS_PHILOSOPHIE					
TERME	**amour**	**beauté**	**connaissance**	**être**	**vérité**
V + T	57	34	11	178	100
T + V	39	54	8	22	0
TOTAL_VERBE	96	88	19	200	100
N + T	51	151	148	114	102
T + N	40	15	90	0	24
TOTAL_NOM	91	166	238	114	126
A + T	0	0	0	0	0
T + A	388	1	278	182	214
TOTAL_ADJECTIF	388	1	278	182	214
TOTAL_TERME	575	255	535	496	440

2.5.5 Aéronautique vs philosophie

Les CLS du domaine de l'aéronautique se répartissent de façon plus uniforme entre les catégories verbales, nominales et adjectivales que pour le domaine de la philosophie, où l'on remarque une nette prédominance des CLS du type T + A.

2.6 Inventaire des cooccurrents

Pour chacun des termes clés, nous avons effectué une première sélection de cooccurrents dans chacune des catégories, soit cinq (5) dans l'une ou l'autre des catégories verbales, verbe + terme ou terme + verbe; cinq (5) dans l'une ou l'autre

des catégories nominales, nom + terme ou terme + nom; et cinq (5) autres dans l'une ou l'autre des combinaisons adjectivales, adjectif + terme ou terme + adjectif, ce qui totalise 15 cooccurrents par terme clé. Dans la base Aéronautique, nous avons relevé un total de 74 combinaisons, et dans la base Philosophie, de 70 combinaisons[23].

Certains critères ont présidé à la sélection des cooccurrents :

1. les cooccurrents ne sont pas synonymes;
2. en présence d'un déverbal, ne pas utiliser le verbe correspondant.

Le tableau V ci-dessous, que nous avons importé du fichier Excel®, dresse la liste des cooccurrents sélectionnés dans la base Aéronautique :

[23] Aucune combinaison n'ayant été relevée dans certaines catégories.

Tableau V : Cooccurrents sélectionnés dans la base Aéronautique

PREMIERE RECHERCHE DE COOCCURRENTS - BASE AERONAUTIQUE											
terme clé	**V + T**		**T + V**		**N + T**		**T + N**		**A + T**		**T + A**
aéronef	autoriser	x	décoller	x	circulation	x	départ	x			certifié
	exploiter	x	franchir	x	contrôle	x					exploité
	maîtriser	x			pilote	x					lourd
					position	x					civil
											IFR
aéroport	atterrir	x			gestion	x			grand	x	contrôlé
	implanter	x			coût	x					international
	desservir	x			contrôleur	x					désigné
	fréquenter	x			activité	x					
	intégrer	x			trafic	x					
piste	dégager	x			circuit	x					transversal
	croiser	x			seuil	x					préférentiel
	poser (se)	x			cap	x					actif
	emprunter	x			angle	x					secondaire
	survoler	x			état	x					sec
vitesse	atteindre	x	augmenter	x			croisière	x	faible	x	minimal
	prescrire	x					sécurité	x			initial
	reprendre	x					montée	x			vraie
	calculer	x					vol	x			normalisé
							descente	x			
vol	régir	x	traverser	x	trajectoire	x	instruction	x			militaire
	planifier	x			plan	x					plané
	maintenir	x			itinéraire	x					rectiligne
	contrôler	x			temps	x					lent
											mixte

Quant au tableau VI ci-dessous, également importé du fichier Excel®, il dresse la liste des cooccurrents sélectionnés dans la base Philosophie :

Tableau VI : Cooccurrents sélectionnés dans la base Philosophie

PREMIERE RECHERCHE DE COOCCURRENTS - BASE PHILOSOPHIE										
terme clé	**V + T**		**T + V**		**N + T**		**T + N**		**A + T**	**T + A**
amour	reconnaître	x	tendre	x	essence	x	soi	x		spirituel
			dominer	x	pureté	x	charité	x		sensible
			unifier	x	mouvement	x				humain
			intégrer	x						interpersonnel
										divin
beauté	retrouver	x	pénétrer	x	contemplation	x	âme	x		intelligible
			recueillir	x	expérience	x	présence	x		
			donner	x	fondement	x				
			susciter	x						
connaissance	conduire	x	procéder	x	principe	x	coeur	x		métaphysique
	évoquer	x			acte	x	chose	x		intuitif
	devancer	x					vrai	x		universel
	procurer	x								intellectuel
										sensible
être	saisir	x	manifester (se)	x	absolu	x				absolu
	justifier	x			achèvement	x				intérieur
	révéler	x			consentement	x				empirique
	penser	x			quête	x				abstrait
					valeur	x				indivis
vérité	atteindre	x			voie	x	intelligence	x		ontologique
	percevoir	x			nature	x				inné
	posséder	x			ordre	x				immuable
	approcher	x								premier
	rencontrer	x								éternel

À partir de ces cooccurrents, nous sommes ensuite retournée dans le corpus pour effectuer une deuxième recherche de combinaisons lexicales afin de voir avec quels termes ces cooccurrents étaient utilisés.

2.7 Deuxième recherche à partir des cooccurrents

Nous avons donc procédé à une deuxième recherche de contextes à l'aide du logiciel Naturel Pro[MC], mais cette fois, à partir des cooccurrents sélectionnés à la première étape décrite plus haut. Par exemple, pour le terme clé *aéronef*, nous avons retenu le cooccurrent *décoller*. Nous sommes donc retournée dans la base de textes pour voir si le verbe *décoller* se combinait avec d'autres termes (comme *avion*, *appareil*, etc.). Nous avons relevé tous ces termes de même que le nombre d'occurrences de chacun d'eux, que nous avons consignés dans un autre fichier Excel® (voir l'annexe 5, dont le tableau VII ci-dessous présente un exemple).

Tableau VII : Deuxième recherche de termes — *autoriser*[24]

T de départ	V + T	cooccurrent	fréquence
aéronef	autoriser	aéronef	29
		calage	2
		conducteur	1
		décollage	9
		fournisseur	2
		installations de communication	1
		minimum	2
		org. de contrôle	1
		personne	2
		pilote	13
		transport	1
		vol	7
			70

[24] L'annexe 5 présente quelques résultats de la deuxième recherche de cooccurrents.

Soulignons que le critère de fréquence (précisé dans la quatrième colonne du tableau VII ci-dessus) constitue un excellent indicateur de la lexicalisation d'une combinaison. En effet, si la fréquence d'une combinaison donnée est très élevée, nous avons de bonnes raisons de croire qu'il s'agit d'une combinaison fortement lexicalisée, et par voie de conséquence, soit d'une collocation, soit d'une unité terminologique propre ou en voie de le devenir. Toutefois, il faut garder à l'esprit qu'au regard de notre corpus relativement limité, la fréquence a pour but principal d'attester de l'occurrence des termes.

Nous nous sommes d'abord assurée, pour chacun des cooccurrents faisant l'objet de la deuxième série de combinaisons, de distinguer :

1. chacun des sens véhiculés par les cooccurrents,
2. la structure syntaxique, qui donne lieu à différents régimes selon qu'il s'agisse, si nous prenons l'exemple des verbes, de combinaisons avec des compléments d'objet directs ou avec des compléments d'objet indirects.

Nous avons ensuite retenu uniquement les combinaisons dont le sens véhiculé par le cooccurrent était le même que celui véhiculé dans la CLS où apparaît le terme clé de départ.

Prenons l'exemple du verbe *autoriser* que nous avons relevé parmi les cooccurrents de *aéronef* dans la base Aéronautique. Ce verbe comporte plusieurs acceptions, dont les deux suivantes :

- «(qqn) à, lui accorder un droit, une permission» (sens (1) dans la figure 20 ci-dessous) (Robert 1995 : 164).

- «permettre, rendre possible (sens (2) dans la figure 20 ci-dessous) (Robert 1995 : 164).

Aux fins de notre recherche, nous avons retenu dans l'exemple présent toutes les combinaisons qui relèvent du premier sens du verbe *autoriser*, mais pas celles qui relèvent du deuxième, puisque le premier sens est celui qui est véhiculé dans la CLS sélectionnée à partir du terme clé *aéronef*. La figure 20 ci-dessous illustre le découpage effectué en fonction des différents sens du cooccurrent.

Figure 20 : Distinction entre les différents sens de *autoriser*

autoriser (1) +	aéronef conducteur fournisseur installation UNICOM organisme de contrôle personne pilote
autoriser (2) +	calage décollage transport vol minimum

Examinons maintenant un autre exemple, celui du nom *principe* que nous avons relevé parmi les cooccurrents du terme clé de départ *connaissance* dans la

base Philosophie. Nous avons distingué les deux acceptions suivantes du mot *principe* :

- «Cause, origine ou élément constituant» (sens (1) dans le tableau VIII ci-dessous) (Robert 1995 : 1780).
- «Les vérités fondamentales sur lesquelles s'appuie tout raisonnement» (sens (2) dans le tableau VIII ci-dessous) (Robert 1995 : 1780).

Tableau VIII : Deuxième recherche de termes — *principe*

T de départ	N + T	cooccurrent	fréquence
connaissance	principe	âme	1
		amour	1
		causalité	11
		chose	4
		connaissance	2
		devenir	1
		discernement	1
		esprit	1
		être	13
		évidence intérieure	1
		existence	4
		expérience spirituelle	1
		finalité	3
		identité	7
		individualité	1
		individuation	5
		intellectus	2
		intelligence	1
		intériorité	1
		intersubjectivité	1
		jugement de valeur	1
		liberté	2
		manifestation	1
		métaphysique	10
		morale	2
		non-contradiction	12
		ontologie	1
		ordre	3
		pensée	5
		philosophie	3

T de départ	N + T	cooccurrent	fréquence
		raison d'être	6
		réalité	1
		réel	1
		référence	1
		sacré	1
		similitude	4
		substance	1
		unité	4
		vérité	2
		vie	3
			126

Aux fins de notre recherche, nous avons retenu dans l'exemple présent toutes les combinaisons qui relèvent du premier sens du nom *principe*, mais pas celles qui relèvent du deuxième, puisque, encore une fois, il s'agit du sens véhiculé dans la CLS qui a été sélectionnée à partir du terme clé *connaissance.* La figure 21 ci-dessous illustre le découpage effectué en fonction des différents sens du cooccurrent.

Figure 21 : Distinction entre les différents sens de *principe*

principe (1) +	âme	+	discernement
	esprit		évidence intérieure
	être		expérience spirituelle
	pensée		intellectus
	substance		intelligence
	amour		intériorité
	chose		morale
	existence		devenir
	finalité		manifestation
	identité		vie
	individualité		jugement
	ordre		liberté
	réalité		métaphysique
	réel		ontologie
	sacré		philosophie
	connaissance		vérité

principe (2) +	causalité similitude unité intersubjectivité	référence non-contradiction raison d'être

Nous avons répété l'exercice pour l'ensemble des cooccurrents sélectionnés en aéronautique et en philosophie. Les cooccurrents devaient tous véhiculer le même sens que dans la CLS dans laquelle apparaissait le terme clé de départ.

2.8 Regroupement des termes par classes conceptuelles

À cette étape-ci de notre étude, nous avons regroupé les termes clés par classes conceptuelles à partir des données recueillies au cours de la deuxième recherche de cooccurrents, et après avoir purgé nos résultats des CLS non pertinentes, comme nous l'avons expliqué au point 2.7.

Considérons la combinaison suivante tirée de notre corpus : *aéronef* + *décoller*.

Comme nous l'avons déterminé au point 1.3, cette combinaison constitue une collocation conceptuelle. En effet, le verbe peut se combiner à d'autres termes appartenant à la même classe conceptuelle que *aéronef* (*avion, hélicoptère,* etc.*)*.

Ainsi, nous avons poursuivi notre recherche en regroupant les termes clés des CLS à l'aide des étiquettes opératoires que nous avions provisoirement établies (voir le point 2.2), et que nous utilisons ci-après à titre de classes conceptuelles (point 2.8.2 ci-après).

2.8.1 Détermination des classes conceptuelles

Tout d'abord, voici ce que nous entendons par classe conceptuelle. Une classe conceptuelle permet de regouper des termes selon, par exemple, leur nature, leurs propriétés ou leur fonction particulière (Sager 1990 : 35).

Considérons la série de termes suivants : *avion*, *aéronef*, *aéroplane*, *hydravion*, *aérodyne*, *dirigeable*, *aérostat*, *planeur*, *montgolfière.* Ces termes font partie de la classe regroupant les appareils capables de se déplacer dans les airs, cette dernière caractéristique dénotant une «fonction» commune aux neuf (9) termes.

Précisons maintenant notre classification et regroupons parmi ces termes ceux qui possèdent les caractéristiques suivantes :

1. la propriété de se déplacer dans les airs, et
2. celle de voler sans l'aide d'un organe propulseur.

On ne retiendra alors que les termes *planeur*, *aérostat*, *montgolfière* et *dirigeable*. Ces termes pourraient partager, par généralisation, un cooccurrent commun comme *lancer + [un planeur/un dirigeable/un aérostat/une montgolfière]* , dont le sens, en revanche, n'est pas vraiment compatible avec celui des termes *avion, aéronef, aéroplane ou hydravion*.

Ainsi, la définition de classes conceptuelles plus ou moins importantes permet de rendre compte de séries de CLS, comme on le verra plus loin.

2.8.2 Élaboration des classes

Afin de faciliter le regroupement des termes clés, nous avons défini, pour chacun des domaines, une série de classes décrites ci-après.

Pour ce qui est du domaine de l'aéronautique, nous avons constitué neuf (9) classes conceptuelles :

1. AV = appareil volant (par ex., *aéronef*);
2. AC = activité (par ex., *vol*);
3. EA = entité animée (par ex., *pilote*);
4. EB = entité abstraite (par ex., *validité*);
5. EM = entité matérielle (par ex., *bagage*);
6. IT = installations propres au trafic aérien (par ex., *aéroport*);
7. LI = lieu (par ex., *ville*);

8. UM = unité de mesure (par ex., *vitesse*);
9. VC = voie de circulation (par ex., *piste*).

Le domaine de la philosophie comprend huit (8) classes conceptuelles, que nous précisons ci-après :

1. AC = activité (par ex., *expérience*);
2. AM = activité morale (par ex., *recueillement*);
3. EB = entité abstraite (par ex., *idée*);
4. EA = entité animée (par ex., *homme*)
5. EI = entité immatérielle (par ex., *être*);
6. EP = entité physique (par ex., *matière*);
7. DI = disposition (par ex., *bienveillance*);
8. PR = principe (par ex., *vérité*).

Ces classes sont très générales, mais elles se sont révélées suffisantes aux fins de notre étude.

2.8.3 Regroupement

Nous avons regroupé par classes conceptuelles, pour chacun des sens et en fonction des structures syntaxiques, les combinaisons relevées de la manière suivante[25] :

Tableau IX : Regroupement à partir du verbe *autoriser*

VERBE	CLASSE	TERME	
autoriser	AV	aéronef	(v + tcod + à + v)
	EM	bagage	(v + tcod + à + v)
	EA	conducteur	(v + tcod + à + v)
		fournisseur	(v + tcod + à + v)
		personne	(v + tcod + à + v)
		pilote	(v + tcod + à + v)
	IT	installations	(v + tcod + à + v)
		organisme	(v + tcod + à + v)

Une fois cette partie cruciale du processus de recherche franchie — notamment la sélection de termes clés, la recherche de contextes, la sélection des combinaisons lexicales spécialisées, l'inventaire des cooccurrents pour chaque terme clé, la deuxième recherche de cooccurrents, la sélection d'une deuxième série de CLS, le regroupement des termes par classes conceptuelles — nous avons poursuivi notre analyse proprement dite des résultats afin d'évaluer le rapport qui

[25] L'annexe 8 présente quelques résultats du regroupement par classes conceptuelles des CLS retenues.

existe entre les collocations lexicales et les collocations conceptuelles dans les deux domaines de spécialité concernés. Nous présentons les résultats au chapitre suivant.

CHAPITRE III : RÉSULTATS

3 Considérations générales

Notre analyse a porté sur la répartition des collocations lexicales par rapport à celle des collocations conceptuelles. Nous avons désigné comme collocations lexicales toutes les combinaisons lexicales qui, dans le cadre de notre corpus, se soumettaient à la définition telle qu'elle est formulée par Heid (1992) (voir les points 1.2 et 1.3). Rappelons de façon sommaire cette définition : les collocations lexicales sont non généralisables, c'est-à-dire qu'un seul mot clé est utilisé avec un cooccurrent donné. Quant aux collocations conceptuelles, elles s'en distinguent par le fait que les mots clés peuvent être généralisés à des ensembles d'unités lexicales plus vastes et sémantiquement apparentées.

Nous avons fondé notre analyse sur une méthode de découpage des réalités véhiculées par les combinaisons, qui nous a permis de distinguer, *grosso modo*, certaines relations sémantiques grâce à la détermination de classes conceptuelles.

Essentiellement, nous avons procédé en regroupant les séries de termes en fonction de grandes classes dénotant, entre autres, l'activité, la disposition, la propriété (abstraite ou concrète), la nature (matérielle ou immatérielle), la finalité (principe fondamental ou autre), la fonction ou la situation dans l'espace (lieu).

Toutes les CLS qui ont pu être regroupées en fonction de classes conceptuelles se retrouvent dans les collocations conceptuelles. Un examen préliminaire révèle qu'elles sont majoritaires dans notre corpus, que ce soit dans le domaine de l'aéronautique ou dans celui de la philosophie, et qu'elles s'y retrouvent sensiblement dans les mêmes proportions, soit les quatre cinquièmes des combinaisons. Les autres combinaisons ont été comptabilisées comme des collocations lexicales. Toutefois, nous verrons qu'elles ne relèvent probablement pas toutes de cette dernière catégorie, pour des raisons que nous examinons plus loin.

Nous présentons d'abord les résultats obtenus dans le domaine de l'aéronautique, puis dans celui de la philosophie.

3.1 Répartition des collocations — domaine de l'aéronautique

Nous avons retenu 560 combinaisons lexicales spécialisées à partir des 74 cooccurrents sélectionnés. De ce nombre, une analyse préliminaire indique que 484 CLS, soit 86 %, sont des collocations conceptuelles (CC) et que 76 CLS, soit 14 %, sont des collocations lexicales (CL).

Nous avons ensuite recensé les collocations lexicales (CL) et les collocations conceptuelles (CC) en fonction des classes que nous avions

préalablement définies, et nous avons obtenu les résultats indiqués dans le tableau X ci-après :

Tableau X : CC vs CL en fonction des classes conceptuelles (aéronautique)

CLASSE	% des CLS	CC	CL
Appareil volant (AV) :	9 % (53)	80 % (42)	20 % (11)
Activité (AC) :	17% (93)	86 % (80)	14 % (13)
Entité animée (EA) :	4 % (21)	71 % (15)	29 % (6)
Entité abstraite (EB) :	22 % (126)	93 % (117)	7 % (9)
Entité matérielle (EM) :	2 % (13)	62 % (8)	38 % (5)
Installations — trafic (IT) :	13 % (71)	89 % (63)	11 % (8)
Lieu (LI) :	7 % (37)	84 % (31)	16 % (6)
Unité de mesure (UM) :	15 % (82)	15 % (82)	0 % (0)
Voie de circulation (VC) :	11 % (64)	72 % (46)	28 % (18)
TOTAL	100 % (560)	86 % (484)	14 % (76)

Nous observons les faits suivants :

- la classe conceptuelle Entité abstraite (EB) donne lieu au plus grand pourcentage de CLS;

- la classe conceptuelle Entité matérielle (EM) donne lieu au plus faible pourcentage de CLS, suivie de près par la classe Entité animée (EA);

- les autres CLS se répartissent pratiquement dans les mêmes proportions au sein des autres classes;

- de manière générale, le rapport CC – CL est d'environ 4 CC pour 1 CL, à l'exception de la classe Entité abstraite (EB), dans laquelle on note un écart important, soit un rapport de près de 20 CC pour 1 CL, et de la classe Entité matérielle (EM), dans laquelle le rapport est de pratiquement 2 CC pour 1 CL. Remarquons également que seule la classe Unité de mesure (UM) ne comporte que des CC.

Nous avons également recensé le nombre de CC et de CL en fonction des catégories grammaticales, et nous avons obtenu les résultats qui sont présentés au tableau XI ci-dessous :

Tableau XI : CC vs CL en fonction des catégories grammaticales (aéronautique)

CATÉGORIE	% des CLS	CC	CL
VERBALE	35 % (195)	88 % (170)	12 % (25)
NOMINALE	26 % (145)	81 % (116)	19 % (29)
ADJECTIVALE	39 % (220)	90 % (198)	10 % (22)
TOTAL	100 % (560)	86 % (484)	14 % (76)

Nous constatons les faits suivants :

- il n'y a pas d'écarts importants entre les différentes catégories; les CLS s'y répartissent dans des proportions relativement semblables;

- le rapport CC – CL est plus élevé dans la catégorie adjectivale, soit de 9 CC pour 1 CL;

- le rapport CC – CL est d'environ 7 CC pour 1 CL dans la catégorie verbale et d'environ 4 CC pour 1 CL dans la catégorie nominale.

3.1.1 Collocations conceptuelles

Une grande majorité des CLS ont donc pu être regroupées en fonction de classes conceptuelles. Nous présentons ci-après quelques exemples de regroupement par classes conceptuelles.[26]

Combinaisons verbales :

- maîtriser → AV → aéronef (v + tcod)
  appareil (v + tcod)
  avion (v + tcod)

- dégager → VC → aire d'atterrissage (v + tcod)
  issue (v + tcod)
  piste (v + tcod)
  servitude (v + tcod)

[26] Rappelons que l'on trouvera à l'annexe 8 quelques résultats du regroupement par classes conceptuelles des CLS retenues dans notre corpus.

- fréquenter → IT → aéroport (v + tcod)
  aérodrome (v + tcod)
  aérogare (v + tcod)
  installation (v + tcod)
- calculer → UM → distance (v + tcod)
  heure d'arrivée (v + tcod)
  longueur de piste (v + tcod)
  vitesse (v + tcod)

Combinaisons nominales :

- pilote → AV → aéronef (n + de + art + t)
  avion (n + de + art + t)
  hélicoptère (n + de + art + t)
  hydravion (n + de + art + t)
- état → IT → aérodrome (n + de + art + t)
  aéroport (n + de + art + t)
  installation (n + de + art + t)
- altitude
- marge
- niveau
- vitesse

  → EB → sécurité (t + de + n)
  (t + de + n)
  (t + de + n)
  (t + de + n)

Combinaisons adjectivales :

- aéroport
- gare

  → IT → international
- piste
- route

  → VC → préférentiel
- faible → UM → altitude
  amplitude
  dimension
  hauteur
  régime
  vitesse
- descente
- vol

  → AC → plané

Nous constatons donc qu'il est possible de regrouper les combinaisons lexicales spécialisées, mais dans les faits, la fréquence indique que l'usage préfère parfois une combinaison plutôt qu'une autre. Reprenons, par exemple, le terme *aéronef*. Dans les associations de ce terme avec le verbe *exploiter* (voir le tableau XII ci-après), nous avons relevé 42 occurrences de la CLS *exploiter* + *aéronef*. Les fréquences pour les autres termes sont beaucoup moins élevées.

Tableau XII : Fréquences observées — *exploiter*

T de départ	V + T	cooccurrent	fréquence
aéronef	exploiter		
		aéronef	42
		giravion	2
		hélicoptère	2

En revanche, si nous examinons les CLS obtenues avec le verbe *maîtriser*, nous constatons que le terme *avion* y est plus souvent associé que le terme *aéronef* (voir le tableau XIII ci-dessous).

Tableau XIII : Fréquences observées — *maîtriser*

T de départ	V + T	cooccurrent	fréquence
aéronef	maîtriser	aéronef	1
		appareil	2
		avion	6

Cette constatation nous amène à rappeler que, quoique deux termes puissent appartenir à une même classe conceptuelle, à la manière de *avion* et

aéronef (appareil pouvant se déplacer dans les airs), leur commutabilité peut être restreinte en LS en fonction des critères liés à l'usage.

3.1.2 Collocations lexicales

L'examen préliminaire nous a permis de recueillir un certain nombre de combinaisons que nous avons recensées comme des collocations lexicales parce qu'un cooccurrent ne s'employait qu'avec un seul terme. Examinons la liste des combinaisons obtenues.

- ϕ Autoriser + [art] aéronef (emploi méronymique)
- Autoriser + [art] bagage
- Maîtriser + [art] direction
- ϕ Maîtriser + [art] tangage
- ϕ Pilote + décoller (emploi méronymique)
- ϕ Route + [prép] départ
- Circulation + [prép + art] voyageurs
- Position + [prép + art] pilote (emploi méronymique)
- Aéronef + certifié
- Équipement + certifié
- Masse + certifiée
- ϕ Vol + exploité
- ϕ Aéronef + civil
- ϕ Contrôleur + IFR
- ϕ Implanter + [art] aéroport
- ϕ Implanter + [art] piste
- Implanter + [art] radar

- Desservir + [art] circulation
- Intégrer + [art] circuit
- Intégrer + [art] cité
- Intégrer + [art] aéronef
- ϕ Gestion + [prép + art] débit
- ϕ Coût + [prép + art] aéroport
- ϕ Activité + [prép + art] aéroport
- ϕ Trafic + [prép + art] route
- Grand + trafic
- ϕ Aéronef + désigné
- ϕ Dégager + vue
- Circuit + [prép] base
- Circuit + [prép] aérodrome
- Cap + [prép] éloignement
- État + [prép + art] avion
- Piste + transversale
- ϕ Aide visuelle + actif
- Radar + secondaire
- Installation + secondaire
- ϕ Piste + sec
- Prescrire + [art] zone
- ϕ Reprendre + [art] vol
- ϕ Reprendre + [art] circuit
- Pilote + [prép] sécurité
- ϕ Route + [prép] vol
- ϕ Préparatif + [prép] vol
- ϕ Route + vraie
- ϕ Pilote + traverser (emploi méronymique)
- ϕ Planifier + [art] route

- ϕ Aéronef + traverser
- ϕ Vol + traverser (emploi méronymique)
- ϕ Pilote + traverser (emploi méronymique)
- ϕ Maintenir + [art] route
- ϕ Contrôler + [art] aéronef
- Contrôler + [art] véhicule
- Itinéraire + [prép + art] vol
- Instruction + [prép + art] altitude
- Économie + mixte
- ϕ Gare + mixte
- Système + mixte
- ϕ Vol + mixte

Après un examen plus approfondi de cette liste, nous avons fait précéder de la puce ϕ plusieurs combinaisons qui ne sont probablement pas des collocations lexicales, puisque, même si nous n'avons relevé qu'une seule occurrence de la combinaison dans notre corpus, notre intuition nous porte à croire que le mot clé peut être généralisé à des ensembles plus vastes d'unités lexicales.

Par exemple, l'emploi du mot clé *maîtriser* dans la combinaison *maîtriser* + *tangage* pourrait être généralisé à un ensemble d'unités lexicales dénotant le mouvement comme *roulis/lacet*.

De la même façon, les termes *pilote*, dans la combinaison *pilote* + *décoller*; *vol* dans la combinaison *vol* + *décoller*; et *vol* dans la combinaison *vol* + *franchir*,

entre autres, sont des emplois méronymiques et, dans les contextes concernés, ils font référence à l'appareil (*avion*, *aéronef*, *hélicoptère*, etc.). Ainsi, dans une première analyse, nous avons attribué l'étiquette d'origine à certains termes (par ex., dans la combinaison *vol* + *franchir*, nous avons attribué l'étiquette Activité au terme *vol*), mais si l'on examine le contexte, la combinaison que nous avons relevée comme étant une CL serait plutôt une CC si nous prenons en compte l'emploi méronymique. En effet, nous devrions alors attribuer l'étiquette Appareil volant au terme *vol* (*vol* + *franchir*, par exemple), étant donné qu'il est pris dans le sens d'*appareil*.

Quant aux autres combinaisons (précédées de la puce !), il faudrait procéder à une étude fondée sur un corpus composé d'un plus grand nombre de textes pour déterminer leur véritable statut (collocation lexicale ou collocation conceptuelle).

3.2 Répartition des collocations — domaine de la philosophie

Nous avons relevé 962 combinaisons lexicales spécialisées à partir des 70 cooccurrents sélectionnés. De ce nombre, une analyse primaire indique que 887 CLS, soit 92 %, sont des collocations conceptuelles (CC) et que 75 CLS, soit 8 %, sont des collocations lexicales (CL).

Nous avons ensuite recensé les collocations lexicales (CL) et les collocations conceptuelles (CC) en fonction des classes que nous avions préalablement définies, et nous avons obtenu les résultats que nous présentons au tableau XIV ci-dessous :

Tableau XIV : CC vs CL en fonction des classes conceptuelles (philosophie)

CLASSE	% des CLS	CC	CL
Activité (AC) :	14 % (134)	92 % (123)	8 % (11)
Activité morale (AM) :	16% (159)	94 % (150)	6 % (9)
Entité animée (EA) :	0,1 % (1)	0 % (0)	,10 % (1)
Entité abstraite (EB) :	35 % (334)	97 % (323)	3 % (11)
Entité immatérielle (EI) :	14 % (135)	93 % (125)	7 % (10)
Entité physique (EP) :	0,5 % (5)	40 % (2)	60 % (3)
Disposition (DI) :	9,4 % (91)	80 % (73)	20 % (18)
Principe (PR) :	11 % (103)	88 % (91)	12 % (12)
TOTAL	100 % (962)	86 % (887)	14 % (75)

Nous observons les faits suivants :

- comme dans le domaine de l'aéronautique, la classe conceptuelle Entité abstraite (EB) donne lieu au plus grand pourcentage de CLS;

- la classe conceptuelle Entité animée (EA) donne lieu au plus faible pourcentage de CLS, suivie de près par la classe Entité physique (EP);

- les autres CLS se répartissent pratiquement dans les mêmes proportions au sein des autres classes, tout comme dans le domaine de l'aéronautique;

- le rapport CC – CL est de plus de 10 CC pour 1 CL dans quatre cas, soit dans les classes Activité (AC), Activité morale (AM), Entité abstraite (EB) et Entité immatérielle (EI);

- le rapport est de 4 CC pour 1 CL dans deux cas, soit dans les classes Disposition (DI) et Principe (PR);

- nous relevons un cas où le rapport est inversé, c'est-à-dire 1 CC pour 1,5 CL dans la classe Entité Physique (EP), où nous retrouvons un plus grand nombre de CL. Toutefois, le nombre de CLS relevées dans cette classe étant négligeable, nous émettons des réserves quant à la validité de ce rapport.

Nous avons également recensé le nombre de CC et de CL en fonction des catégories grammaticales, et nous avons obtenu les résultats qui sont présentés au tableau XV ci-dessous :

Tableau XV : CC vs CL en fonction des catégories grammaticales (philosophie)

CATÉGORIE	% des CLS	CC	CL
VERBALE	24 % (232)	88 % (204)	12 % (28)
NOMINALE	28 % (280)	90 % (252)	10 % (28)
ADJECTIVALE	48 % (450)	96 % (431)	4 % (19)
TOTAL	100 % (962)	92 % (887)	8 % (75)

Nous constatons les faits suivants :

- nous remarquons un écart notable entre les différentes catégories; la catégorie adjectivale comprend deux fois plus de CLS que les catégories verbale et nominale, lesquelles regroupent quant à elles sensiblement le même pourcentage de CLS;

- le rapport CC – CL est plus élevé dans la catégorie adjectivale, tout comme dans le domaine de l'aéronautique, soit de plus de 10 CC pour 1 CL;

- le rapport CC – CL est sensiblement le même dans les catégories verbale et nominale, soit d'environ 7 CC pour 1 CL dans la catégorie verbale, et de 9 CC pour 1 CL dans la catégorie nominale.

3.2.1 Collocations conceptuelles

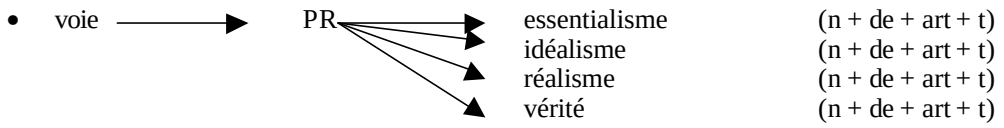
Tout comme dans le domaine de l'aéronautique, une grande majorité des CLS du domaine de la philosophie ont pu également être regroupées en fonction de classes conceptuelles. Nous présentons ci-après quelques exemples.

Combinaisons verbales :

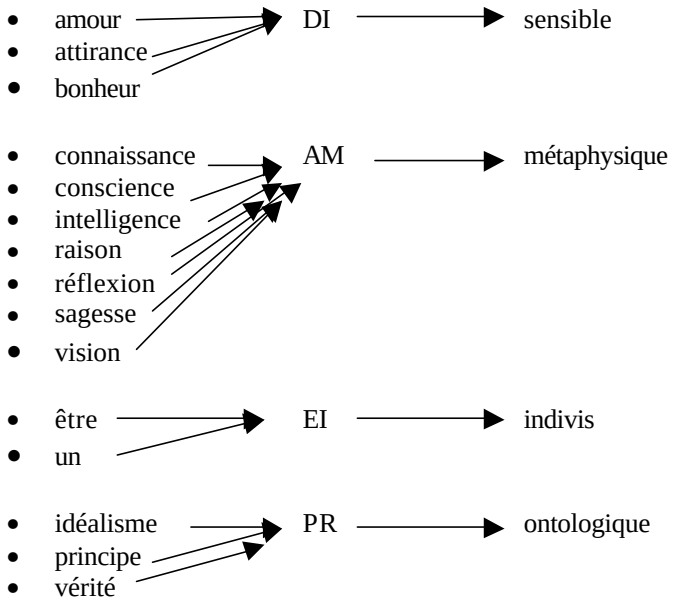
• reconnaître	PR	amour	(v + tcod)
		pureté	
• beauté • bien	EB	susciter	(ts + v)
• conduire	AM	connaissance	(v + tcoi)
		conscience	
• saisir	EI	âme	(v + tcod)
		Dieu	(v + tcod)
		esprit	(v + tcod)
		essence	(v + tcod)
		être	(v + tcod)
		moi	(v + tcod)
		pensée	(v + tcod)
• atteindre	PR	absolu	(v + tcod)
		quiddité	(v + tcod)
		universalité	(v + tcod)
		vérité	(v + tcod)

Combinaisons nominales :

• expérience	EB	beauté	(n + prép + t)
		réel	(n + prép + t)
		sacré	(n + prép + t)
• acte	AM	connaissance	(n + de + t)
		intellection	(n + de + t)
		intelligence	(n + de + t)
		réflexion	(n + de + t)



Combinaisons adjectivales :



3.2.2 Collocations lexicales

Comme au point 3.1.2, nous présentons les combinaisons qui ont été classées, à la suite d'un premier examen, parmi les collocations lexicales. Une étude subséquente nous a également amenée à mettre en doute certaines de ces collocations, que nous avons également fait précéder de la puce ϕ.

ϕ Amour + tendre

- Amour + dominer
- Expérience + dominer
- Amour + unifier

- Beauté + unifier
- Esprit + unifier
- Sagesse + unifier
- Amour + intégrer
- ϕ Essence + [prép + art] amour
- Essence + [prép + art] métaphysique
- Pureté + [prép + art] acte
- ϕ Pureté + [prép + art] amour
- Mouvement + [prép + art] amour
- Mouvement + [prép + art] appétition
- Homme + sensible
- Matière + sensible
- Corps + humain
- Métaphysique + humain
- Amour + interpersonnel
- ϕ Vérité + divin
- Amour + pénétrer
- Beauté + pénétrer
- Conscience + pénétrer
- Expérience + pénétrer
- Beauté + recueillir
- Intériorité + recueillir
- Valeur + recueillir
- ϕ Contemplation + [prép] Dieu
- Expérience + [prép + art] accueil
- ϕ Fondement + [prép + art] amour
- Philosophie + [prép + art] âme
- Beauté + [prép + art] présence
- Discernement + [prép + art] présence

- Pureté + [prép + art] présence
- Beauté + intelligible
- Conduire + [art] âme
- Devancer + [art] connaissance
- ϕ Phénomène + procéder
- ϕ Principe + [prép + art] amour
- Acte + amour
- Acte + [prép] consentement
- Lumière + [prép + art] + coeur
- Mouvement + [prép + art] coeur
- ϕ Connaissance + [prép + art] vrai
- Philosophie + [prép + art] vrai
- Bienveillance + universelle
- Saisir + [art] perfection
- Justifier + [art] donnée
- Justifier + [art] réalisme
- Révéler + [art] vérité
- Penser + [art] grâce
- ϕ Intériorité + manifester (se)
- Absolu + [prép + art] être
- ϕ Quête + [prép + art] vérité
- Quête + [prép + art] bien
- Séparation + [prép + art] absolu
- Homme + intérieur
- ϕ Analyse + empirique
- ϕ Analyse + abstraite
- ϕ Atteindre + [art] réalisation
- ϕ Percevoir + [art] vérité
- Approcher + [art] vérité

- Rencontrer + [art] vérité
- Voie + [prép + art] nature
- Voie + [prép + art] être
- Voie + [prép + art] piété
- Nature + [prép + art] affirmation
- Ordre + [prép] charité
- Intelligence + innée
- Vérité + immuable
- Impuissance + immuable
- Défiance + première
- Substance + première
- Substance + éternelle

3.3 Comparaisons entre les deux domaines

Remarquons d'abord que les CC et les CL se répartissent également dans les deux domaines, soit 86 % de CC en aéronautique et en philosophie; et 14 % de CL en aéronautique et en philosophie.

Pour ce qui est de la répartition des CLS dans les différentes catégories, nous remarquons un plus grand écart entre les catégories dans le domaine de la philosophie, soit 48 % de CLS dans la catégorie adjectivale, 28 % dans la catégorie nominale et 24 % dans la catégorie verbale. Tandis que dans le domaine de l'aéronautique, les CLS se répartissent de façon plus uniforme dans les différentes catégories, soit 39 % dans la catégorie adjectivale, 35 % dans la

catégorie verbale et un pourcentage légèrement moins élevé dans la catégorie nominale, soit 26 %.

En ce qui concerne la répartition des CLS en fonction des diverses classes conceptuelles, nous soulignons une majorité de CLS regroupées dans la classe Entité abstraite (EB) en philosophie, soit 35 %; de même qu'en aéronautique, soit 22 %.

Les CC sont nettement majoritaires dans les deux domaines, mais nous notons une exception dans l'une des classes conceptuelles du domaine de la philosophie, soit 60 % des CL contre 40 % des CC regroupées dans la classe Entité physique (EP). Cependant, nous réitérons les réserves que nous avons formulées au point 3.2 quant à ces résultats.

CONCLUSION

De manière générale, nous avons dégagé certains faits :

1. il est tout à fait possible de regrouper les CLS en fonction de classes conceptuelles auxquelles appartiennent les termes, ce qui donne lieu à des collocations conceptuelles (CC);

2. les collocations conceptuelles (CC) sont majoritaires dans les deux domaines concernés, mais on y trouve également une quantité non négligeable de collocations lexicales (CL) (dont le cooccurrent ne s'emploie qu'avec un seul terme);

3. les comportements lexicaux que nous avons vérifiés en évaluant le rapport entre les CC et les CL relevées dans notre corpus appuient les «intuitions terminologiques», si nous pouvons nous exprimer ainsi, des auteurs étudiés dans le cadre de notre mémoire. En effet, notre recherche montre la pertinence de tenir compte d'une approche conceptuelle dans le traitement terminologique des CLS, nos résultats démontrant à tout le moins l'importance des CC dans les deux domaines concernés (rappelons ces résultats : en aéronautique, le rapport CC – CL varie de 4 CC pour 1 CL à 9 CC pour 1CL; tandis qu'en philosophie, le rapport varie de 7 CC pour 1 CL à plus de 10 CC pour 1 CL).

À la lumière d'un deuxième survol de nos résultats, nous avons également constaté que le nombre de collocations lexicales, tant en aéronautique qu'en philosophie, est probablement moins élevé. Nous émettons toutefois des réserves quant à l'utilisation privilégiée d'un mot clé (ou base) dans notre corpus. En effet, le fait de n'avoir relevé qu'une seule occurrence d'une CLS donnée dans le corpus ne signifie pas pour autant qu'il s'agit là d'une CL. Il demeure qu'en regard de plusieurs CLS que nous avons marquées comme étant des CL, la généralisation semblait possible.

Nous pensons qu'un vaste échantillonnage est souhaitable pour cerner les tendances de comportement applicables au plus grand nombre de cas possibles et dégager les différentes cooccurrences. Notre corpus étant limité en raison des restrictions liées à la production d'un mémoire de maîtrise, nous désirons souligner que les résultats obtenus ne reflètent pas nécessairement le comportement des combinaisons relevées dans tous les domaines de spécialité. Nous croyons cependant qu'ils montrent une tendance générale en LS.

En outre, nous aurions pu porter notre choix sur des domaines d'étude d'un niveau technique encore plus «pointu» afin de mieux faire ressortir les différences de comportement des combinaisons lexicales. Néanmoins, nous avons préféré nous en tenir à des domaines qui nous étaient suffisamment familiers de sorte que notre propre compétence linguistique puisse nous permettre de discriminer

aisément les unités terminologiques proprement dites des autres combinaisons à l'étude.

De plus, un corpus limité à deux domaines de spécialité permet uniquement de faire certaines observations, mais il faudrait étendre la recherche à plusieurs autres domaines pour pouvoir les vérifier. Les chercheurs disposent aujourd'hui d'outils informatisés extrêmement utiles. Nous nous permettons de regretter cependant que ces outils, vu leur énorme potentiel, ne soient pas suffisamment exploités.

Nous souhaitons avant tout que notre mémoire permette de faire avancer l'étude des combinaisons lexicales spécialisées. Nous espérons également qu'il ouvre la voie à d'autres recherches visant à élaborer des règles régissant la combinatoire lexicale ou à fournir aux professionnels de la langue des outils terminologiques plus complets.

Enfin, il est permis d'espérer que l'explosion de la technologie de l'informatique, du multimédia et des moyens de communication permettra de mettre à profit la richesse de la combinatoire terminologique et de créer des outils qui facilitent l'accès à des données terminologiques dépassant la limite du terme unique.

BIBLIOGRAPHIE

BALLY, C. (1944) : *Linguistique générale et linguistique française*, Berne, A. Francke S. A., Benteli S. A.

BENSON, M., BENSON, E. et R. ILSON (1986) : *The BBI Combinatory Dictionary of English: A Guide to Word Combinations*, Amsterdam/Philadelphia, John Benjamins.

BLAIS, E. (1993) : «Le phraséologisme. Une hypothèse de travail», *Terminologies nouvelles* n° 10, pp. 50-56.

CLAS, A. (1994) : «Collocations et langues de spécialité», *Meta*, vol. 39, n° 4, pp. 576-580.

COHEN, B. (1983) : «Lexique de cooccurrents — Bourse et conjoncture économique», mémoire de maîtrise, Montréal, Université de Montréal.

COHEN, B. (1986) : *Lexique des cooccurrents — Bourse et conjoncture économique*, Montréal, Linguatech.

CRYSTAL, D. (1980) : *Dictionary of Linguistics and Phonetics*, 4e édition, Oxford, New York, Blackwell Publishers.

DUBOIS, J. *et al.* (1994) : *Dictionnaire de linguistique et des sciences du langage*, Paris, Larousse.

FONTENELLE, T. (1996) : «Réseaux sémantiques et dictionnaires bilingues électroniques», *Lexicomatique et dictionnairiques*, A. Clas, P. Thoiron et H. Béjoint (dir.), Beyrouth, FMA, pp. 31-42.

GROSS, G. (1988) : «Degré de figement des noms composés», *Langue française*, n° 90, pp. 57-71.

HAUSSMANN, F.J. (1979) : «Un dictionnaire des collocations est-il possible?», *Travaux de linguistique et de littérature* XVII (1), pp. 187-195.

HEID, U. (1992) : «Décrire les collocations. Deux approches lexicographiques et leur application dans un outil informatisé», *Terminologie et traduction*, n° 2-3, pp. 523-548.

HEID, U. (1994) : «On the Way Words Work Together - Topics in Lexical Combinatorics», Martin, W. *et al. Euralex '94 Proceedings*, Amsterdam, The Congress, pp.226-257.

HEID, U. et G. FREIBOTT (1991) : «Collocations dans une base de données terminologique et lexicale», *Meta*, vol. 36, n° 1, mars, pp. 77-91.

KJAER, A.L. (1990) : «Phraseology Research – The State of the Art : Methods of Describing Word Combinations in Language for Specific Purposes», *IITF Journal* 1 (1-2), pp. 3-20.

KOCOUREK, R. (1982) : *La langue française de la technique et de la science*, Brandstetter Verlag, Wiesbaden.

L'HOMME, M. C. (1993) : «Les unités phraséologiques verbales et leur représentation en terminographie», *Terminologie et traduction*, n° 2-3, pp. 493-503.

L'HOMME M. C. (1995) : «Processing Word Combinations in Existing Termbanks», *Terminology*, vol. 2, n° 1, pp. 141-162.

L'HOMME, M. C. (1998) : «Caractérisation des combinaisons lexicales spécialisées par rapport aux collocations de langue générale», Fontenelle *et al*. *Actes Euralex 1998*, Liège, Université de Liège, pp. 513-522.

L'HOMME M. C. et R. GEMME (1997) : «Modèle d'accès informatisé aux combinaisons lexicales spécialisées : verbe + nom (terme) et extension aux nom (déverbal) + préposition : nom (terme), Lapierre, L., I. Oore et H.R. Runte, *Mélanges de linguistique offerts à Rostislav Kocourek*, Université Dalhousie (Halifax, Canada) : Les Presses ALFA, pp. 89-103.

LAINÉ, C., PAVEL, S. et M. BOILEAU (1992) : «La phraséologie - Nouvelle dimension de la recherche terminologique», Travaux du module canadien du Rint, *L'Actualité terminologique* 25 (3), pp. 5-9.

LAPORTE, I. (1996) : *Méthode de recensement et de présentation des combinaisons lexicales en langue de spécialité, accompagnée d'un micro-lexique combinatoire de la pharmacologie cardiovasculaire*, Travail dirigé, Montréal, Université de Montréal.

LAPORTE, I. (1997) : «Recensement et consignation des combinaisons lexicales en langue de spécialité : Exemple d'application dans le domaine de la pharmacologie cardiovasculaire», *Terminologies nouvelles*, n° 16, pp. 95-101.

MARTIN, W. (1992) : «Remarks on Collocations in Sublanguages», dans *Terminologie et traduction* 2-3, pp. 157-164.

MEL' UK, I. *et al*. (1984) : *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I,* Montréal, Les Presses de l'Université de Montréal.

MEL' UK, I. *et al*. (1988) : *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques II,* Montréal, Les Presses de l'Université de Montréal.

MEL' UK, I. *et al*. (1992) : *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III,* Montréal, Les Presses de l'Université de Montréal.

MEL' UK, I., CLAS, A. et A. POLGUÈRE (1995) : *Introduction à la lexicologie explicative et combinatoire (ILEC)*, Louvain-la-Neuve (Belgique), Éditions Duculot, AUPELF UREF, 256 p.

MEL' UK, I et L. WANNER (1996) : «Lexical Functions and Lexical Inheritance for Emotion Lexemes in German», L, Wanner (Ed.), *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam / Philadelphia, John Benjamins, pp. 207-277.

MEYER, I. et K. MACKINTOSH (1994) : «Phraseme Analysis and Concept Analysis in Exploring a Symbiotic Relationship in the Specialized Lexicon», Martin W. *et al.* 1994. *Euralex '94 Proceedings*, Amsterdam, pp. 339-348.

MEYER, I. et K. MACKINTOSH (1996) : «Refining the terminographer's concept-analysis methods: How can phraseology help?», *Terminology* 3 (1), pp. 1-26.

MINISTÈRE DU SECRÉTARIAT D'ÉTAT (1996) : *Termium; la banque de données linguistiques du gouvernement du Canada*, Bureau de la traduction, gouvernement du Canada.

OLF (1996) : *Le DOC de l'Office de la langue française*, Les publications du Québec, gouvernement du Québec.

PAVEL, S. (1993 a) : «Vers une méthode de recherche phraséologique en langue de spécialité», *L'actualité terminologique* 26 (2), pp. 9-13.

PAVEL, S. (1993 b) : «La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques», *Terminologies nouvelles*, n° 10, pp. 23-35.

PAVEL, S. (1994) : *Guide de la recherche phraséologique en langue de spécialité*, Bureau de la traduction, pagination discontinue.

ROBERT, P (1995) : *Le nouveau petit Robert*, Paris, Dictionnaires Le Robert.

SAGER, J. C. (1990) : *A Practical Course in Terminology Processing*, Amsterdam/Philadelphia, John Benjamins Publishing Company.

ANNEXE 1 : OUVRAGES PORTANT SUR L'AÉRONAUTIQUE

CAUSSE, S. *et al.* (1991) : *Phraséologie, approche non radar/par la Subdivision Circulation aérienne, Département exploitation*, 2e édition, Toulouse, École nationale de l'aviation civile.

COT, P. D. A. (1963) : *Les aéroports*, Paris, Presses universitaires de France, coll. Que sais-je?, no 1048.

DANGLETERRE D. (1992) : *L'anglais pour voler*; *Classement alphabétique et thématique du vocabulaire aéronautique pour les pilotes et les contrôleurs*, 2e édition, édité et diffusé par l'auteur.

MINISTÈRE DU TRANSPORT (sans date) : *Manuel de pilotage des avions et Guide de l'instructeur*, Ottawa, Direction de l'aviation civile.

TRANSPORTS CANADA (1981) : *Manuel de vol aux instruments*, 2e édition, Ottawa, Ministre des Approvisionnements et Services Canada.

TRANSPORTS CANADA (1995) : *A.I.P. Canada; Publication d'information aéronautique*, Ottawa, Services d'information aéronautique, pagination discontinue.

TRANSPORTS CANADA (1995) : *Procédure aux instruments*, Géomatique Canada, ministère des Ressources naturelles, pagination discontinue.

ANNEXE 2 : OUVRAGES PORTANT SUR LA PHILOSOPHIE

COLIN, A. (1956) : *Nouveau vocabulaire philosophique*, Paris, Armand Cuvillier.

D'AQUIN THOMAS (1964) : *L'être et l'esprit*, Paris, Presses universitaires de France.

DESCARTES, R. (1963) : *Oeuvres philosophiques : Tome I*, Paris, Garnier.

DUROZOI, G. et A. ROUSSEL (1990) : *Dictionnaire de philosophie*, Paris, Nathan.

GARDEIL, H. D. (1960) : *Initiation à la philosophie de S. Thomas D'Aquin : Métaphysique*, Paris, Éditions du Cerf.

GEIGER L. B. (1963) : *Philosophie et spiritualité*, Paris, Éditions du Cerf.

HEIDEGGER, M. (1967) : *Introduction à la métaphysique*, Paris, Éd. Gallimard.

JOLIVET, R. (1962) : *Vocabulaire de la philosophie*, Lyon, Emmanuel Vitte.

LEGRAND, G. (1983) : *Dictionnaire de philosophie*, Paris, Éd. Bordas.

MASSET P. (1989) : *L'intériorité retrouvée : la philosophie spirituelle d'Aimé Forest*, Paris, Tequi.

ANNEXE 3 : LISTE DE COOCCURRENTS DU DOMAINE DE L'AÉRONAUTIQUE

ANNEXE 4 : LISTE DE COOCCURRENTS DU DOMAINE DE LA PHILOSOPHIE

ANNEXE 5 : DEUXIÈME RECHERCHE DE COOCCURRENTS

ANNEXE 6 : RÉSULTATS — AÉRONAUTIQUE

ANNEXE 7 : RÉSULTATS — PHILOSOPHIE

ANNEXE 8 : REGROUPEMENT PAR CLASSES CONCEPTUELLES

ANNEXE 9 : EXTRAIT DE NOTRE BASE DE DONNÉES (CORPUS)

ANNEXE 10 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES VERBALE

ANNEXE 11 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES NOMINALE

ANNEXE 12 : EXEMPLE D'UNE PAGE DE LA BASE DE DONNÉES ADJECTIVALE