L’expression de quatre fonctions lexicales syntagmatiques dans le français parlé

par

Julia Pugachenkova
BA in French, Saint-Petersburg State University, 2019

A Project Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ART

in the Department of French

©Julia Pugachenkova, 2020
University of Victoria

All rights reserved. This project may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Comité de supervision

L'expression de quatre fonctions lexicales syntagmatiques dans le français parlé

par

Julia Pugachenkova
BA in French, Saint-Petersburg State University, 2019

Comité de supervision

Dr. Catherine Caws (Department of French)
Superviseuse

Dr. Catherine Léger, (Department of French)
Seconde lectrice

Résumé

Ce projet de recherche propose une analyse de l'expression de quatre fonctions lexicales syntagmatiques dans les français parlé. La notion des fonctions lexicales a été abordée par plusieurs chercheurs. Cependant, selon mes recherches, on trouve peu d'études qui examinent les termes utilisés pour exprimer des fonctions lexicales concrètes et qui donnent un panorama de ces termes employés dans différents registres et différentes variétés de la langue française. Cette étude, tout en s'inscrivant dans le cadre de la théorie linguistique Sens-Texte, vise à mettre en lumière les adjectifs représentant les fonctions lexicales Bon, AntiBon, Magn et AntiMagn employés dans des collocations du type « adjectif + nom » et « nom + adjectif » au sein du français parlé, ainsi que leur fréquence et les différences quant aux choix des adjectifs selon les variables propres aux locuteurs (telles que sexe, âge, origine et lieu de résidence).

Table de matières

Comité de supervision ii
Résumé iii
Table de matières iv
Liste des tableaux v
Liste des figures vi
Remerciements vii
1. Introduction 1
2. État présent de la recherche et problématique 3
2.1. La théorie linguistique Sens-Texte 3
2.2. La définition du terme *collocation* 6
2.3. La notion des fonctions lexicales 15
2.4. Problématique de la recherche 17
3. Méthodologie 18
3.1. Description du corpus 18
3.2. Méthode utilisée 21
4. Analyse des résultats 25
4.1. Résultats globaux 25
4.2. Adjectifs utilisés pour former des collocations 25
4.2.1. Adjectifs dans les collocations illustrant la FL Bon 26
4.2.2. Adjectifs dans les collocations illustrant la FL AntiBon 27
4.2.3. Adjectifs dans les collocations illustrant la FL Magn 28
4.2.4. Adjectifs dans les collocations illustrant la FL AntiMagn 29
4.3. La fréquence des adjectifs 30
4.4. Choix des adjectifs selon les variables propres aux locuteurs 32
4.4.1. Sexe des locuteurs 32
4.4.2. Âge des locuteurs 34
4.4.3. Origine des locuteurs 35
4.4.4. Lieu de résidence des locuteurs 37
5. Conclusion 39
Bibliographie 43
Annexe 1. Toutes les collocations extraites en contexte selon les FL. 47

Liste des tableaux

Tableau 1. Trois classes majeures de phrasèmes. 9
Tableau 2. Âge des locuteurs. 20
Tableau 3. Origine des locuteurs. 20
Tableau 4. Algorithme de l'encodage des FL. 24
Tableau 5. Fréquence des adjectifs utilisés dans les collocations extraites. 31
Tableau 6. Répartition des trois adjectifs selon le sexe des locuteurs. 33

Liste des figures

Figure 1. Thèmes de la conversation dans le corpus FrancoToile. 19

Remerciements

Tout d'abord, je tiens à remercier ma superviseuse Catherine Caws qui m'a fait connaître la théorie Sens-Texte et qui m'a guidée inlassablement tout au long de ma recherche. J'apprécie énormément tous les conseils que vous m'avez donnés dans les moments d'incertitude.

Je suis aussi très reconnaissante à Catherine Léger, ma seconde lectrice. Merci de nous avoir donné tellement d'outils indispensables dans la recherche. Je me souviendrai toujours de vos mots qu'il faut être un peu stressé pour progresser.

Et un grand merci à mes parents. Je vous dois tout ce que j'ai dans la vie.

1. Introduction

Le présent projet est consacré à l'expression des fonctions lexicales Bon, AntiBon, Magn et AntiMagn dans le français parlé. Cette étude, inscrite dans le cadre de la théorie linguistique Sens-Texte, porte sur les collocations du type « adjectif + nom » et « nom + adjectif » au sein du français parlé. Bien que dans le cadre de la théorie Sens-Texte il existe un nombre important de recherches consacrées à la description lexicographique des associations lexicales et à leur représentation dans les dictionnaires (voir, par exemple, Mel'čuk, 2003), on trouve peu d'études qui appliquent cette approche à la linguistique de corpus, surtout à des corpus de la langue parlée. En outre, peu d'études examinent les termes concrets utilisés pour former des syntagmes qui illustrent des fonctions lexicales décrites dans le cadre de la théorie Sens-Texte. De plus, on voit un manque de recherches qui donnent un panorama de ces termes employés dans différents registres et différentes variétés de la langue française. Ainsi, en ayant recours aux notions des fonctions lexicales, cette étude propose une nouvelle perspective du traitement et de l'analyse des collocations formées par les locuteurs de français dans leurs énoncés oraux spontanés.

Cette recherche est basée sur l'approche théorique Sens-Texte dont les principes les plus saillants sont présentés dans la section 2, aussi bien que deux notions clé du projet : celles des termes *collocation* et *fonctions lexicales*.

La section 3 vise à décrire le corpus et la méthode utilisés. Ma méthodologie est fondée sur l'étude d'Alonso Ramos (2000) qui a élaboré des critères pour l'encodage des collocations représentant des fonctions lexicales différentes.

La section 4 est consacrée à l'analyse des résultats obtenus. J'y présente les éléments de réponse aux trois questions de recherche suivantes : Quels adjectifs les locuteurs du français utilisent-ils dans leurs énoncés oraux pour former des collocations représentant les fonctions lexicales Bon, AntiBon, Magn et AntiMagn ? Quelle est la fréquence des adjectifs utilisés pour former les collocations en question ? Observe-t-on des différences quant au choix des adjectifs pour former ces collocations selon l'origine géographique des locuteurs, ainsi que selon d'autres variables, comme, par exemple, l'âge des locuteurs ?

2. État présent de la recherche et problématique

Cette partie du travail vise à présenter les trois axes théoriques sur lesquels mon étude est basée, ainsi que les questions de recherche qui guideront mon analyse. Premièrement, dans la section 2.1, je décris les principes fondateurs de la théorie linguistique Sens-Texte qui servira de cadre pour mon étude. Deuxièmement, dans la section 2.2, je me tourne vers le débat mené autour de la notion du terme *collocation* afin de choisir la définition sur laquelle je m'appuierai pour parler de l'objet de ma recherche. De plus, je mentionne les manières de traiter les collocations suggérées par plusieurs chercheurs. Troisièmement, la section 2.3 est consacrée à une autre notion qui est aussi essentielle pour cette étude : celle des fonctions lexicales. Finalement, dans la section 2.4, je présente mes questions de recherche auxquelles j'essayerai de fournir des éléments de réponse dans le cadre de ce projet.

2.1. La théorie linguistique Sens-Texte

La théorie Sens-Texte trouve son origine dans les années 1960 quand cette nouvelle approche a été mise en avant par un groupe de linguistes soviétiques. Depuis l'avènement de cette théorie, ses principes fondamentaux ont été décrits dans plusieurs ouvrages et articles (entre autres Kahane, 2003 ; Mel'čuk, 1974, 1981, 1985, 1988, 1993, 1997 ; Polguère, 1998, 2011). Dans l'aperçu de la théorie Sens-Texte qui suit, je m'appuierai principalement sur la retranscription de la leçon inaugurale donnée par I. Mel'čuk au Collège de France (1997) qui, à mon avis, est la source qui présente le cadre théorique en question de la manière la plus complète.

Selon Mel'čuk (1997), la théorie Sens-Texte[1] est fondée sur trois postulats. En premier lieu, cette approche considère toute langue naturelle comme « une correspondance multi-multivoque entre un ensemble dénombrable infini de sens et un ensemble dénombrable infini de textes » (p. 4). Ici, le terme *sens* doit être conçu dans son acception la plus étroite possible, c'est-à-dire comme un « invariant des paraphrases langagières » (Mel'čuk, 1993, p. 42). Quant au terme *texte*, il se rapporte dans ce cas-là à chaque représentation physique du sens donné et non seulement aux textes cohérents dont le discours est bien organisé.

En deuxième lieu, la TST suggère que cette correspondance entre les sens et les textes doit être représentée sous forme d'un modèle fonctionnel de la langue naturelle dont l'élaboration est une des fins principales de cette théorie. Ces modèles fonctionnels sont fondés sur le passage du sens vers le texte (Sens => Texte), car la méthodologie de la TST consiste à étudier la synthèse des énoncés (ou bien à décrire ce processus) et non pas à les analyser. La place centrale est donc accordée à la production de la parole, tandis que le décodage des énoncés est considéré comme une activité plutôt secondaire. En d'autres mots, la TST est une approche onomasiologique qui modélise avant tout l'activité langagière du locuteur et non pas celle du destinataire. Cela signifie que la TST ne s'intéresse pas au discours au sens dialogique, mais pourtant les énoncés dont la synthèse est décrite dans le cadre de cette théorie peuvent être rapprochés au terme *parole* pris au sens saussurien. Il est à noter que la TST donne de l'importance à la synonymie des énoncés, ce qui veut dire que les modèles fonctionnels créés au sein de cette approche

[1] Dorénavant la TST.

sont basés sur le paraphrasage. Ainsi, la capacité du locuteur d'exprimer le même sens à travers plusieurs textes doit être prise en compte.

Le troisième postulat de la TST découle du deuxième et concerne la description de la correspondance entre les sens et les textes. Pour créer un modèle fonctionnel complet, on a besoin de deux niveaux intermédiaires entre le sens (la représentation sémantique) et le texte (ou bien la représentation physique du sens). Le rôle de ces niveaux est joué par la représentation morphologique et la représentation syntaxique qui décrivent les régularités spécifiques du mot et de la phrase respectivement.

Dans cet ordre d'idées, on peut résumer la TST en cinq caractéristiques qui ont été relevées par Polguère (1998). Premièrement, cette théorie considère toute langue comme un outil virtuel qui permet aux locuteurs de traduire des sens en énoncés, c'est-à-dire, en textes. Cela signifie que la TST s'oppose aux approches théoriques du type génerativiste et s'intéresse aux énoncés réels produits plutôt qu'aux possibilités d'énoncés. Deuxièmement, cette théorie est fondée sur des principes généraux qui peuvent être appliqués à toutes les langues existantes. Troisièmement, tout en reposant sur des principes universels, cette approche permet de construire des modèles pour chaque langue humaine spécifique. Quatrièmement, ces modèles construits au sein de la TST sont calculables et de nature computationnelle. Cinquièmement, la théorie en question fait appel aux langages formels pour représenter les énoncés linguistiques en prenant en compte tous les niveaux du fonctionnement de la langue (la phonétique, la morphologie, la sémantique, la syntaxe).

Comme mentionné ci-dessus et comme le remarque Mel'čuk (1993), « l'objectif de la théorie Sens-Texte est d'assurer le système de référence et les formalismes

nécessaires à l'écriture des modèles Sens-Texte » (p. 47). Cependant, les buts de la TST ne se bornent pas à la construction des modèles fonctionnels. Cette approche vise aussi l'élaboration du « dictionnaire idéalisé » (Mel'čuk, Clas, et Polguère, 1995, p. 29) ou bien du Dictionnaire Explicatif et Combinatoire[2] (Mel'çuk, 1999). Selon Mel'čuk, Clas, et Polguère (1995), la description d'un lexème dans le DEC doit inclure cinq zones : zone sémantique, zone syntaxique, zone de combinatoire lexicale, zone d'exemples et zone phraséologique. En proposant de consacrer deux zones de chaque entrée du DEC (zone de combinatoire lexicale et zone phraséologique) aux associations lexicales, les auteurs soulignent, semble-t-il, l'importance de ce phénomène. Ainsi, on peut conclure que, dans le cadre de la TST, une place assez considérable est accordée aux unités phraséologiques, y compris les collocations. C'est la raison pour laquelle j'ai décidé de baser mon étude, dont l'objet est la collocation, sur cette théorie. Pourtant, il faut remarquer que le terme *collocation* et sa définition posent souvent un problème pour les chercheurs. Dans la section suivante, j'essaye d'esquisser des points de vue de différents chercheurs sur cette notion.

2.2. La définition du terme *collocation*

Déjà en 1957, Firth (p. 179) affirmait : « On peut connaître le mot selon sa compagnie »[3]. Cette phrase reste, semble-t-il, extrêmement pertinente dans le domaine de la linguistique de nos jours car une place assez importante est accordée à la notion de collocation et aux manières d'analyser ce phénomène.

---

[2] Le DEC en abrégé.

[3] Ma traduction.

Parmi l'abondance des recherches consacrées à la phraséologie et au figement lexical (entre autres Benigno, Grossmann et Kraif, 2015 ; Bolly, 2010 ; Gries, 2013 ; Mel'čuk, 2003, 2012, 2013), une certaine divergence de points de vue existe quant à la définition du terme *collocation*. Cette hétérogénéité a été mise en avant par Mel'čuk (2013), par Tutin et Grossmann (2002) et par Gries (2013). Toutefois, ce dernier a évoqué quelques caractéristiques qui ont tendance à apparaître de façon récurrente dans les définitions données au terme *collocation* par plusieurs chercheurs. Premièrement selon l'analyse des définitions menée par Gries, les chercheurs ont tendance à considérer comme éléments constitutifs des collocations des mots et non pas des catégories grammaticales ou des parties de discours. Gries note que si en décrivant les éléments constitutifs d'une association lexicale les chercheurs ont recours à des catégories plus larges, comme, par exemple, des parties de discours, ils utilisent déjà les termes *colligation* ou *collostruction* et non pas *collocation*. Deuxièmement, on définit les collocations le plus souvent comme des associations polylexicales, c'est-à-dire qu'elles contiennent au moins deux mots. Troisièmement, le nombre d'occurrences d'une association de mots qui permet de la définir comme une collocation n'est pas d'habitude établi de manière stricte. Quatrièmement, les chercheurs acceptent que les collocations soient caractérisées par une certaine flexibilité syntaxique et que, tout en étant liés syntaxiquement, les éléments des collocations puissent ne pas être directement adjacents. Cinquièmement, on affirme souvent que, du point de vue sémantique, le choix des éléments des collocations est en quelque sorte imprédictible ou même arbitraire. Cette thèse semble être dans une certaine mesure soutenue par Legallois et Tutin (2013). Ces chercheurs suggèrent que la caractéristique déterminante des unités phraséologiques est

celle du figement lexical qui est principalement défini à travers les critères suivants : la compositionnalité/la non-compositionnalité sémantique, l'actualisation/la non-actualisation des référents, le degré de la fixité syntaxique et de la restriction combinatoire. Les auteurs illustrent ces critères en citant l'exemple suivant qui est basé sur une expression figée :

> Pour une expression comme *accorder ses violons*, par exemple, on notera ainsi (a) que le sens de l'expression ne se construit pas de façon régulière à partir des composants, (b) que le référent de violons n'est pas accessible, (c) que l'expression ne se prête guère aux alternances syntaxiques (**ses violons ont été accordés* ; *les violons qu'ils ont accordés*), (d) que les composants de l'expression ne commutent pas facilement avec des synonymes ou des co-hyponymes (**accorder les violoncelle*s ; **ajuster ses violons*). (p. 4)

En effet, tous les critères décrits ci-dessus sont plus ou moins utilisés dans les définitions du terme *collocation* proposées par plusieurs auteurs (voir entre autres Bolly, 2010 ; Gries 2013 ; Mel'čuk, 2013). Par exemple, Bolly (2010) définit les collocations comme « des séquences polylexicales dont les constituants entretiennent une relation lexicalement contrainte, tout en conservant leur caractère compositionnel et leur contenu catégoriel propres » (p. 15). Cela signifie qu'il s'agit de cooccurrences lexicales restreintes dont les constituants coapparaissent d'une manière plutôt arbitraire. Selon

Bolly, les collocations sont composées de bases et de collocatifs, ceux-ci étant sélectionnés par celles-là.

Mel'čuk (2013), en parlant de la composition des collocations, distingue aussi les bases de la collocation des collocatifs de celle-ci. Mel'čuk décrit les bases de la collocation comme des « composantes sélectionnées librement par le locuteur » et note que des collocatifs doivent être choisis « en fonction du sens à exprimer » et en fonction de la base (p. 135). Par exemple, dans l'association lexicale *porter une accusation* la base *accusation* nécessite le collocatif *porter*, tandis que dans *faire allusion* la base *allusion* est combinée avec le collocatif *faire*. Cependant, cet auteur considère les collocations comme une classe majeure des phrasèmes (celles-là étant un sous-ensemble de ceux-ci) qui, à leur tour, sont définis comme des énoncés multilexémiques[4] non libres. Mel'čuk propose toute une typologie de phrasèmes qui peut être résumée comme dans le tableau suivant :

Compositionnalité des phrasèmes / Nature des contraintes	non compositionnels	compositionnels
lexicales	LOCUTIONS (*manger la grenouille*[5])	COLLOCATIONS (*porter une accusation*)
sémantico-lexicales	**impossible**	CLICHÉS (*sauf imprévu*)

Tableau 1. Trois classes majeures de phrasèmes.

[4] Terme employé par Mel'čuk, synonyme de *polylexicales*.

[5] Tous les exemples cités dans ce tableau sont empruntés à Mel'čuk.

Ainsi, Mel'čuk postule que les collocations sont des phrasèmes lexicaux compositionnels et semi-contraints. Cette dernière caractéristique se rapporte au fait que la base de la collocation est choisie librement par le locuteur, tandis que le collocatif est choisi en fonction de la base.

En ce qui concerne la compositionnalité sémantique des phrasèmes, pour montrer comment elle doit être conçue, Mel'čuk (2013) propose la formule suivante :

> Un signe linguistique complexe AB est dit compositionnel ssi[6]
> AB = A ⊕ B. Le symbole « ⊕ » représente l'opération
> d'union linguistique, qui réunit les signes et leurs composantes
> selon leur nature et leurs propriétés suivant les règles générales
> de la langue donnée. (p. 131)

Cependant, Mel'čuk (2013) souligne que cette opposition, « la compositionnalité *vs* la non compositionnalité des phrasèmes », n'est pas à confondre avec l'opposition « la transparence *vs* l'opacité des associations lexicales », car la première dichotomie est strictement binaire et tout à fait objective, tandis que la deuxième peut avoir des degrés et est plutôt subjective, parce qu'elle dépend de la capacité mentale du locuteur de discerner la métaphore cachée derrière l'association lexicale donnée aussi bien que de son niveau langagier. Par exemple, le phrasème *châteaux en Espagne* est certainement non compositionnel, puisque son signifié ʿprojet irréalisableʾ ne se compose pas des signifiés des parties constitutives du phrasème en question. Par contre, son degré d'opacité (qui est, sans doute, lié à sa non-compositionnalité) ne peut être défini que d'une manière

[6] Si et seulement si. Abréviation employée par Mel'čuk.

subjective. Quant aux phrasèmes compositionnels, on peut citer l'exemple de *café noir* dont le signifié est composé des signifiés de *café* et de *noir*.

Quant à la nature des contraintes phraséologiques qui consiste en la sélection des composants des phrasèmes, Mel'čuk (2013) suggère qu'il faut distinguer les phrasèmes lexicaux des phrasèmes sémantico-lexicaux. Les phrasèmes lexicaux ne dépendent pas de la situation de la parole et d'autres facteurs extralinguistiques. En revanche, les phrasèmes sémantico-lexicaux sont construits autour de la situation de la parole concrète. Ainsi, ce sont des phrasèmes récurrents qu'on emploie fréquemment comme tels. D'après Mel'čuk, les phrasèmes sémantico-lexicaux constituent une des trois classes majeures qu'il appelle les clichés. Les phrasèmes *Vous dites ?* et *Défense de stationner* peuvent servir d'exemples de clichés parmi lesquels on trouve aussi des formules de politesse.

Pour décrire le degré de flexibilité syntaxique propre aux collocations, Mel'čuk (2013) propose les notions de collocations standard et de collocations non standard. Les collocations standard, selon Mel'čuk, peuvent subir le paraphrasage[7] lexico-syntaxique, comme, par exemple, en (1).

1. Le public a vivement applaudi son discours. = Le public a couvert son discours d'applaudissements. (p. 136)

Les collocations non standard, en revanche, « manifeste[nt], entre la base et le collocatif, un lien sémantique non systématique : ce lien s'applique à très peu de bases, très souvent à une seule base ; et il n'implique que très peu de collocatifs, le plus souvent

[7] Terme emprunté à Mel'čuk.

un seul » (Mel'čuk, 2013, p. 136). *Café noir*, *nuit blanche*, *année bissextile* sont des exemples de collocations non standard. Parmi ces trois collocations, c'est *année bissextile* qui est la plus illustrative quant au lien sémantique non systématique, car l'adjectif *bissextile* ne se combine d'habitude qu'avec le nom *année*. On peut aussi remarquer que *café noir* est moins métaphorique que *nuit blanche*. La métaphore de *nuit blanche* ne sera pas gardée si on traduit ce phrasème vers d'autres langues. Ainsi, on peut considérer cette hétérogénéité comme un défaut de la classification des phrasèmes proposée par Mel'čuk.

Quant à Polguère (2003b), qui travaille aussi dans le cadre de la TST, il vise à définir les collocations en partant d'une autre notion, notamment de celle d'idiomaticité, et les présente comme des expressions semi-idiomatiques :

> L'expression AB (ou BA), formée des lexies A et B, est une collocation si, pour produire cette expression, le locuteur sélectionne A librement d'après son sens ʿAʾ, alors qu'il sélectionne B pour exprimer un sens ʿCʾ en fonction de A. (p. 134)

Néanmoins, dans sa définition, Polguère (2003b) fait aussi appel aux notions de la base de la collocation et du collocatif, en mentionnant que celle-là contrôle celui-ci.

Les notions de bases et de collocatifs apparaissent aussi chez Tutin et Grossmann (2002) qui, pourtant, proposent une typologie différente de celle mise en avant par Mel'čuk (2013). Ces auteurs suggèrent que :

> Une collocation est l'association d'une lexie (mot simple ou phrasème) L et d'un constituant C (généralement une lexie, mais parfois un syntagme par exemple *à couper au couteau*

dans *un brouillard à couper au couteau*) entretenant une relation syntaxique telle que :

- C (le collocatif) est sélectionné en production pour exprimer un sens donné en cooccurrence avec L (la base) ;
- Le sens de L est habituel. (p. 12)

Dans leur typologie, les auteurs identifient les collocations opaques (arbitraires et non transparentes sur le plan sémantique), les collocations transparentes (dont le collocatif est difficilement prédictible, mais facilement interprétable, comme, par exemple *de loup* dans l'expression *faim de loup*) et les collocations régulières dont les composants sont associés d'une façon motivée et transparente (par exemple, *nez aquilin*, *grande tristesse*, etc.).

Dans cet ordre d'idées, on peut dire que les définitions de *collocation* proposées par ces chercheurs ont plusieurs points communs. Dans le cadre de mon étude, je m'appuierai sur la définition mel'čukienne du terme *collocation* basée sur une typologie assez détaillée ce qui, à mon avis, permet de mieux distinguer les collocations des autres phénomènes phraséologiques. De plus, afin de distinguer les collocations des autres types de phrasèmes, Mel'čuk (2013) a recours aux critères de la compositionnalité / la non-compositionnalité et non pas à celles de l'opacité / la transparence des associations lexicales comme le font Tutin et Grossmann (2002). Les critères privilégiés par Mel'čuk, qui se base sur le fait que le signifié de l'association lexicale est ou n'est pas composé des signifiés de ses parties constitutives, semblent être plus objectifs. Je ne traiterai donc que des collocations standard qui sont marquées d'une certaine flexibilité syntaxique et dont les collocatifs peuvent s'associer à des bases différentes (par exemple : un livre

passionnant = un livre qui est passionnant, le jour le plus passionnant, un métier passionnant, un sujet passionnant, etc.).

En ce qui concerne la façon dont on aborde les collocations dans les études lexicologiques, on peut y relever deux tendances principales décrites par Bolly (2010). Il s'agit de l'approche fonctionnelle, représentée notamment par Granger et Paquot (cités par Bolly, 2010), et de l'approche statistique. Cette dernière s'intéresse à toutes les unités lexicales dont les éléments ont une attraction lexico-grammaticale, même minimale, et utilise largement des critères de récurrence et de cooccurrence. Quant à l'approche fonctionnelle, elle met l'accent sur le figement sémantique des unités lexicales et, contrairement à l'approche statistique, ne considère pas, par exemple, la cooccurrence *le frère* comme une collocation, car elle n'est pas revêtue d'une signification compositionnelle, puisque *le* dans ce cas ne peut pas être considéré comme un élément qui soit doté d'une sémantique à part dans cette association de mots.

Le fait que la linguistique d'aujourd'hui est vraiment caractérisée par ces deux tendances relativement à la notion de collocation est également discuté par Benigno, Grossmann et Kraif. (2015). Ces auteurs suggèrent aussi que ce sont les approches mixtes, combinant les critères statistiques et l'analyse sémantique, qui permettent de traiter des collocations de la manière la plus objective et qui deviennent de plus en plus répandues parmi les chercheurs. Pourtant, les deux approches décrites ci-dessus se rapportent plutôt, semble-t-il, à l'aspect définitoire de collocation puisque les chercheurs qui les adoptent cherchent principalement à comprendre ce qui peut être désigné par le terme *collocation*. De plus, quand il s'agit du phénomène collocationnel dans le cadre de

la TST, on ne peut pas laisser de côté la notion des fonctions lexicales à laquelle la section suivante est consacrée.

2.3. La notion des fonctions lexicales

Les études menées au sein de la TST qui traitent des collocations et d'autres types d'associations lexicales font souvent appel aux fonctions lexicales[8] (voir, par exemple, Diachenko, 2006 ; L'Homme, 2002 ; Mel'čuk, 1982, 1998 ; Polguère, 2003a, 2003b). Bien que les définitions du concept des FL proposées par tous ces chercheurs soient presque identiques, je voudrais me référer à celle qui a été mise en avant par Polguère (2003b), car on la trouve dans une des sources les plus récentes consacrées aux FL.

La FL est définie par Polguère (2003b) de la manière suivante :

> Une ***fonction lexicale f*** décrit une relation existant entre une lexie **L**—l'argument de f — et un ensemble de lexies ou d'expressions figées appelé la valeur de l'application de f à la lexie L. La fonction lexicale f est telle que :
>
> 1. l'expression f (L) représente l'application de f à la lexie L ;
> 2. chaque élément de la valeur de f (L) est lié à L de la même façon.
>
> Il existe autant de fonctions lexicales qu'il existe de types de liens lexicaux et chaque fonction lexicale est identifiée par un nom particulier : Syn, Anti, etc. (p. 131)

[8] FL en abrégé.

Pour une meilleure compréhension de la notion des FL, Polguère (2003a) propose de les considérer comme des métalexies qui « sont caractérisées par des propriétés de sens et de combinatoire, mais ne sont pas associées de façon directe à des signes, et donc à des signifiants » (p. 126). Selon Polguère, c'est au locuteur d'appliquer ces métalexies à des lexies particulières en tant que fonctions.

Polguère (2003b) ainsi que Mel'čuk, Clas, et Polguère (1995) distinguent les FL paradigmatiques, qui expriment les relations entre lexies sur la base des rapports sémantiques (à savoir synonymie, antonymie, etc.), et les FL syntagmatiques, qui expriment toutes sortes de relations lexicales entre lexies au sein d'une collocation. Comme l'objet de mon étude est la collocation, je me concentrerai sur les FL syntagmatiques.

Parmi les configurations des FL les plus courantes, Polguère (2003a) cite les fonctions Bon, AntiBon, Magn et AntiMagn. Selon Polguère (2003b), la FL Bon « associe à une lexie l'ensemble des lexies ou expressions linguistiques qui expriment auprès d'elle le sens général 'bon', 'bien'... — c'est-à-dire marquent l'évaluation positive ou l'approbation du locuteur » (p. 137), tandis que AntiBon sert à exprimer le sens de l'évaluation négative ou la désapprobation du locuteur. Ainsi, la collocation *défaite honteuse* peut servir d'exemple de la FL AntiBon. La collocation *frisson agréable* illustre, par contre, la FL Bon. Quant aux FL Magn et AntiMagn, celle-là « associe à une lexie l'ensemble des lexies ou expressions linguistiques qui expriment auprès d'elle (en tant que modificateurs) l'intensification, c'est-à-dire le sens général 'intense', 'très', 'beaucoup', etc. » (p. 137), tandis que celle-ci est le contraire d'intensificateur. Dans cet ordre d'idées, la collocation *un grand frisson* illustre la FL Magn et la collocation *un*

frisson léger illustre la FL AntiMagn. Comme la TST utilise largement les notations mathématiques pour encoder et formaliser les phénomènes langagiers, on peut présenter les collocations données ci-dessus de la manière suivante : AntiBon(défaite) = honteuse ; Bon(frisson) = agréable ; Magn(frisson) = grand ; AntiMagn(frisson) = léger.

Toute les FL citées ci-dessus sont des FL adjectivales et, par conséquent, ont tendance à être illustrées par des collocations du type « adjectif+nom » et « nom+adjectif » qui sont l'objet de la présente recherche.

2.4. Problématique de la recherche

Bien que la notion des FL soit largement éclairée dans la littérature, à ma connaissance, cet instrument lexicographique n'a pas été appliqué à des corpus oraux. Cependant, cette démarche pourrait permettre de mieux comprendre la complexité et l'hétérogénéité du français dans sa forme orale ainsi que de tracer les tendances d'utilisation du lexique dans la langue courante. Ainsi, vu le manque de travaux portant précisément sur les moyens d'expression et de manifestation des FL particulières dans le français parlé, mes questions de recherche sont les suivantes :

1. Quels adjectifs les locuteurs du français utilisent-ils dans leurs énoncés oraux pour former des collocations représentant les FL Bon, AntiBon, Magn et AntiMagn ?
2. Quelle est la fréquence des adjectifs utilisés pour former les collocations en question ?
3. Observe-t-on des différences quant au choix des adjectifs pour former ces collocations selon l'origine géographique des locuteurs, ainsi que selon d'autres variables, comme, par exemple, l'âge des locuteurs ?

3. Méthodologie

Cette partie du projet est consacrée à la présentation du corpus qui m'a servi de source des données à analyser (section 3.1), ainsi qu'à la description de la méthode utilisée dans le cadre de la recherche (section 3.2).

3.1. Description du corpus

Dans le cadre de cette recherche, je me concentre sur des énoncés oraux et sur des différences hypothétiques qui peuvent exister dans le choix des adjectifs dans les collocations du type « adjectif + nom » et « nom + adjectif » qui illustrent les FL Bon, AntiBon, Magn et AntiMagn. Comme source des données, j'ai décidé d'utiliser le corpus FrancoToile. Ce corpus est constitué d'énoncés oraux de locuteurs francophones qui ont été recueillis entre 2008 et 2014. L'objectif principal du corpus FrancoToile consistait à recueillir les points de vue des locuteurs de français sur la francophonie et la culture francophone dans le monde entier.

Les énoncés sont présentés sous forme de 83 vidéos dont la durée totale est de 3 h 17 m 44 s. La durée moyenne d'une vidéo est de 2 m 22 s. La durée inférieure de la vidéo est de 49 secondes, tandis que la durée supérieure de la vidéo est de 5:22 minutes. Bien que les énoncés du corpus FrancoToile aient été recueillis sous forme de vidéos, j'extrais les collocations qui sont l'objet de mon étude en me basant sur les transcriptions des vidéos. Ainsi, le nombre total de mots du corpus est de 31 393 mots et le nombre de mots moyen par vidéo est de 369 mots. Il est à noter que les énoncés recueillis dans le corpus FrancoToile ne sont pas préparés à l'avance, mais cependant les locuteurs sont conscients du fait que leurs propos sont enregistrés, et de plus, le thème de la conversation fait partie d'un choix proposé par l'intervieweur et non sélectionné complètement librement par le

locuteur. Toutefois, les thèmes proposés sont variés et vont de la francophonie en milieu minoritaire jusqu'aux vacances ou aux loisirs. La répartition des thèmes dans les vidéos est présentée dans le graphe suivant :

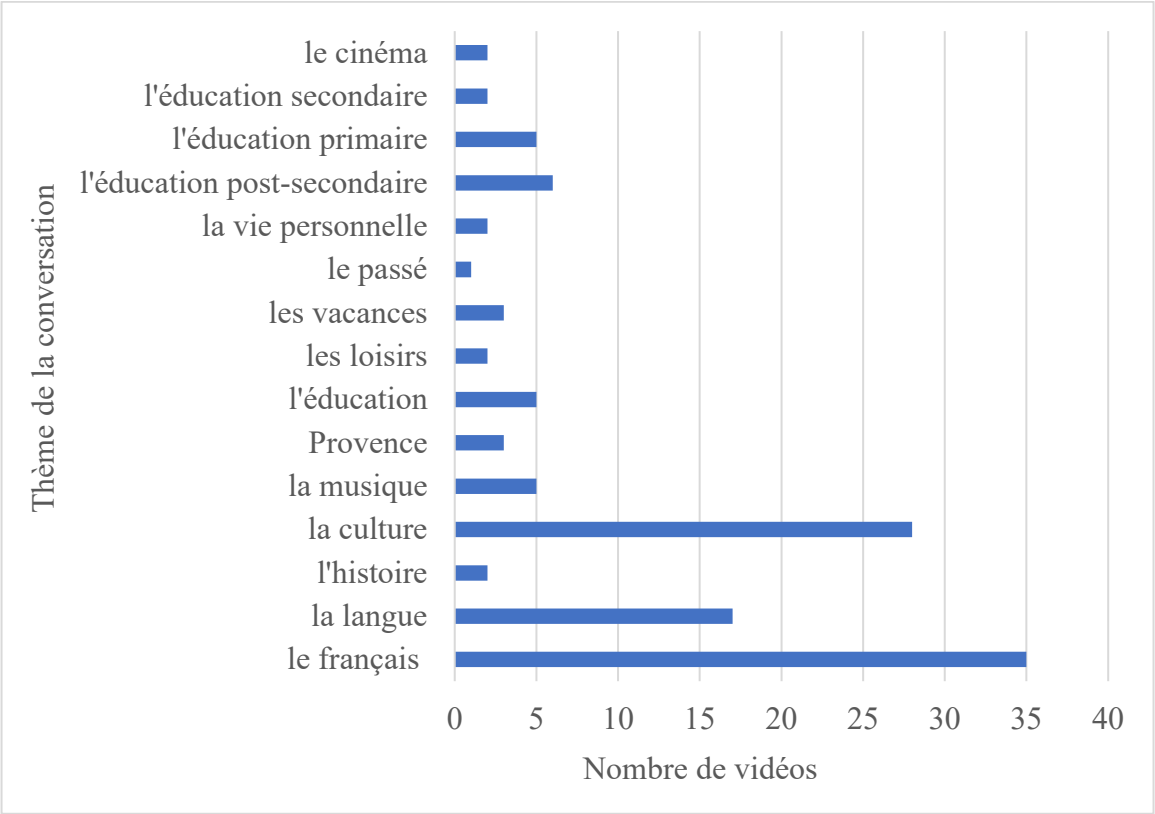

Figure 1. Thèmes de la conversation dans le corpus FrancoToile.

Dans le cadre du projet FrancoToile, on a enregistré les énoncés de 49 locuteurs dont 23 hommes et 26 femmes. Les locuteurs sont d'âges variés, allant de jeunes enfants aux personnes âgées. On peut donc relever huit tranches d'âge qui sont présentées dans le tableau suivant :

Tranche d'âge	Nombre de locuteurs	Pourcentage
de 0 à 9 ans	3	6,12 %
de 10 à 14 ans	10	20,4 %
de 15 à 18 ans	1	2,04 %
de 19 à 25 ans	17	34,69 %
de 26 à 39 ans	8	16,32 %
de 40 à 54 ans	3	6,12 %
de 55 à 64 ans	4	8,16 %
65 ans +	3	6,12 %

Tableau 2. Âge des locuteurs.

Ces locuteurs proviennent de différentes régions du Canada et de la France, mais aussi de la Suisse, de l'île Maurice, du Cameroun et d'Haïti. La répartition des locuteurs selon leur pays d'origine est illustrée par le tableau suivant :

Pays d'origine	Nombre de locuteurs	Pourcentage
Cameroun	1	1,96 %
Canada	20	39,21 %
France	24	47,05 %
Haïti	1	1,96 %
Île Maurice	2	3,92 %
Suisse	1	1,96 %

Tableau 3. Origine des locuteurs.

Ce qui est aussi important dans le corpus FrancoToile c'est le lieu de résidence des locuteurs au moment de l'enregistrement qui peut m'aider à identifier les conséquences que l'environnement linguistique peut avoir sur leurs choix langagiers. Au

moment de l'enregistrement, les locuteurs résidaient au Rwanda, au Burkina Faso, au Mali, en Californie, à Île Maurice, dans des régions et des villes différentes de la France (Ardèche, L'île d'Yeu, La-Roche-Sur-Yon, Nantes, Paris, Rennes, Saint-Rémy-de-Provence, Strasbourg, Pays de la Loire, Ouest de la France, Avignon) et du Canada (la ville de Québec, Moncton, Montréal, Nord du Québec, Nord-est du Québec, Ottawa, Victoria). Il est à noter que le corpus ne me donne pas d'informations concernant le statut social des locuteurs, telles que leur niveau d'éducation, leur profession, etc.

3.2. Méthode utilisée

Comme l'objectif principal de cette étude consiste en une identification des adjectifs qui font partie des collocations illustrant les FL Bon, AntiBon, Magn et AntiMagn (de type « adjectif + nom » et « nom + adjectif ») dans les énoncés oraux des locuteurs, il est nécessaire de choisir un système qui me permettra de relever les collocations en question. Un tel algorithme accompagné de critères heuristiques pour l'encodage des collocations sont suggérés par Alonso Ramos (2000). Dans la présentation de la méthode utilisée qui suit, je me baserai largement sur la procédure élaborée par cette chercheuse.

Les FL placées au centre de mon étude sont des FL syntagmatiques. Cela signifie que, tout d'abord, il me faut choisir un ou des critères afin de distinguer les collocations illustrant des FL syntagmatiques de celles qui représentent des FL paradigmatiques. Alonso Ramos (2000) propose quatre critères clairs qui permettent d'identifier la nature des collocations. Selon cette chercheuse, la FL représentée par la collocation est syntagmatique si le mot-clé (ou la base de la collocation) et la valeur (ou bien le collocatif) sont cooccurrents (p. 464). De plus, la nature des relations sémantiques

existant entre la base et le collocatif peut être prise en compte (p.465). Ainsi, la dérivation sémantique indiquera une FL paradigmatique, tandis que la relation sémantique du type « prédicat-argument » est un trait caractéristique d'une FL syntagmatique. Ensuite, parmi les collocations représentant des FL syntagmatiques il faut choisir celles du type « nom + adjectif » ou « adjectif + nom ». En outre, comme dans ce projet il s'agit de collocations standard qui sont assez flexibles du point de vue syntaxique et qui peuvent subir le paraphrasage, il est nécessaire de prendre en considération les collocations dont les éléments ne sont pas adjacents (comme, par exemple, *le livre passionnant* extrait de la phrase *Le livre que je lis est passionnant*). Finalement, en me basant sur le sens du mot-clé et de la valeur, je peux identifier la FL concrète qui est illustrée par chaque collocation. Ainsi, les collocations dont le collocatif exprime les sens ʿ très ʾ, ʿ intense ʾ, ʿ à un degré élevé ʾ ou les sens opposés illustrent les FL Magn ou AntiMagn respectivement. Quant aux FL Bon et AntiBon, les collocations qui les représentent exprimeront l'approbation ou la désapprobation subjective du locuteur.

Sous cette optique, on peut présenter l'algorithme élaboré ci-dessus dans le tableau suivant :

Étape	Critère(s)	Exemples
1. Relever des collocations illustrant les FL syntagmatiques.	- Le mot-clé et la valeur sont cooccurrents ; - le mot-clé et la valeur entretiennent une relation sémantique de type « prédicat-argument ».	Les lexies *crime* et *criminel* ne sont pas cooccurrents et représentent la dérivation sémantique → une FL paradigmatique. Dans *alimentation équilibrée*, les lexies cooccurrent et la valeur *équilibrée* joue le rôle de l'argument du prédicat *alimentation* → une FL syntagmatique.[9]
2. Choisir des collocations illustrant les FL adjectivales.	- La collocation est composée d'un nom (N) et d'un adjectif (A); - dans les phrases complexes, les composants des collocations ne doivent pas être	*un livre passionnant*, *un livre qui est passionnant*, etc.

[9] Les deux exemples sont empruntés à Alonso Ramos (2000, p. 465).

	directement adjacents.	
3. Identifier les FL concrètes.	Le sens du mot-clé et de la valeur.	*un grand frisson* → Magn *un frisson léger* → AntiMagn *un frisson agréable* → Bon *une défaite honteuse* [10]→ AntiBon

Tableau 4. Algorithme de l'encodage des FL.

Pour le moment, à ma connaissance, il n'existe pas d'outil informatique qui permette l'encodage automatique ou semi-automatique des FL. Ainsi, toute cette procédure est à faire manuellement. Cependant, pour obtenir une classification et représentation plus adéquates des données, j'ai eu recours au logiciel NVivo qui est un outil informatique de traitement des données en analyse qualitative. Ce logiciel permet d'organiser, d'analyser et de synthétiser des données non structurées comme, par exemple, les interviews ou les transcriptions d'énoncés oraux. De plus, NVivo propose de nombreuses options pour représenter les résultats d'analyse d'une manière claire et illustrative sous forme de tableaux et de diagrammes. Grâce à NVivo, je peux identifier la fréquence des adjectifs étudiés et c'est à l'aide de cet outil que je peux établir des liens directs entre le sexe, l'âge, l'origine, le lieu de résidence actuel des locuteurs et leur choix des adjectifs. Finalement, en me basant sur la classification des données extraites, je vais tenter de répondre aux questions de recherche posées dans la partie précédente du travail.

[10] Exemples extraits de Polguère (2003b, p. 138).

4. Analyse des résultats

Cette partie du travail vise à présenter les résultats de l'analyse les plus saillants ainsi que les conclusions auxquelles on peut aboutir compte tenu des résultats obtenus. Chaque section ci-dessous est consacrée à une des questions de recherche présentées dans la partie 2.4 du projet.

4.1. Résultats globaux

Au cours de l'analyse du corpus FrancoToile j'ai extrait 155 collocations du type « adjectif + nom » et « nom + adjectif », y compris celles dont les éléments ne sont pas directement adjacents. Parmi ces 155 collocations, 50 collocations illustrent la FL Bon (32,26 %), 8 collocations illustrant la FL AntiBon (5,16 %), 57 collocations représentent la FL Magn (36,77 %) et 40 collocations (25,80 %) représentent la FL AntiMagn. On peut remarquer qu'il y a une différence notable entre le nombre de collocations extraites qui représentent la FL AntiBon et le nombre de collocations illustrant les autres FL étudiées. Cela peut être expliqué par le fait que les thèmes de la conversation proposés par l'interviewer (tels que la culture francophone, l'éducation ou les vacances) suscitent principalement des émotions positives chez les locuteurs et ne leur font presque pas exprimer la désapprobation.

4.2. Adjectifs utilisés pour former des collocations

Tout d'abord, je m'intéressais aux adjectifs que les locuteurs choisissent pour former des collocations représentant les FL étudiées. Plus bas, je présente des éléments de réponse à cette question de recherche pour chacune des quatre FL.

4.2.1. Adjectifs dans les collocations illustrant la FL Bon

50 collocations représentant la FL Bon sont formées avec les 17 différents adjectifs suivants : *adorable*, *apprécié*, *beau*, *blanc*, *bon*, *enrichissant*, *extraordinaire*, *génial*, *grand*, *haut*, *intéressant*, *joli*, *magnifique*, *positif*, *prisé*[11], *sympa*, *unique*.

Comme mentionné dans la section 2.3, les collocations illustrant la FL Bon expriment l'approbation ou l'évaluation positive du locuteur. La plupart des collocations extraites respectent parfaitement ce critère : *une matière qui est très intéressante*, *de très bons chanteurs*, *une expérience très enrichissante*, etc. Cependant, j'ai également relevé des collocations qui expriment non seulement l'appréciation subjective du locuteur, mais aussi l'approbation qui est propre à un certain groupe dont le locuteur fait partie et dont il partage l'opinion. C'est le cas des collocations formées avec les mots *prisé* et *apprécié*, par exemple, dans *les fêtes folkloriques sont très appréciées* et *chanteurs en langue provençale qui sont aussi très prisés*.

J'ai extrait aussi une collocation qui peut être vue comme une métaphore : *des choses qui étaient blanches* (qui s'oppose à *des choses qui étaient noires* dans *Par exemple pour définir ce qui est bien ce qui est mal / alors que ça on l'apprenait quand même à l'école avant / il y avait des choses qui étaient blanches / il y avait des choses qui étaient noires*). Bien que ce soient des associations lexicales en quelque sorte métaphoriques, ce sont des phrasèmes compositionnels dont le sens est rendu transparent par le contexte, et, par conséquent, selon la typologie de Mel'čuk (2013), on peut les considérer comme des collocations et pas comme des locutions.

[11] Les participes passés qui jouent le rôle de l'adjectif ont aussi été pris en compte.

De plus, il y a trois adjectifs utilisés dans les collocations relevées qui, pris comme tels, n'expriment pas le sens 'bon', 'bien'. Ce sont les adjectifs *unique*, *grand*, *haut*. Toutefois, leur sens dépend, semble-t-il, du nom auquel ils sont associés et on peut constater que les collocations *très hautes études*, *interprétations très uniques* et *nos grandes chanteuses* représentent vraiment la FL Bon.

24 collocations illustrant la FL Bon (telles que *des chansons qui sont géniales*, *le paysage est magnifique*, *les gens sont adorables*, etc.) peuvent être retrouvées dans la section « Cooccurrences » du dictionnaire Antidote[12]. J'ai pu faire quelques observations sur des collocations qui ne sont pas inclues dans le dictionnaire. Selon ce dictionnaire, l'adjectif *beau* s'utilise avec le nom *image*, mais quand il s'agit d'apparence, tandis que dans le corpus étudié, ce nom désigne la représentation mentale (*mais ils ont une très belle image en général de la France*). Le dictionnaire indique que l'adjectif *sympa* est souvent employé avec des noms désignant des êtres humains (*mec sympa*, *gars sympa*, *fille sympa*, *voisin sympa*), ce qui permet de considérer la collocation *les dames de la cantine sont sympas* comme une cooccurrence assez typique pour le français. C'est aussi le cas de l'adjectif *prisé* qui, d'après le dictionnaire, s'utilise avec le nom *conférencier* et qui est combiné avec le nom *chanteur* dans le corpus étudié (*chanteurs en langue provençale qui sont aussi très prisés*).

4.2.2. Adjectifs dans les collocations illustrant la FL AntiBon

Les adjectifs utilisés pour former les 8 collocations illustrant la FL AntiBon sont *bête*, *épouvantable*, *mauvais*, *noir*, *pollué*. Dans ce cas-ci, on peut aussi noter qu'en

[12] Les collocations qui ne sont pas indiquées dans le dictionnaire sont mises en caractères gras dans Annexe 1.

choisissant les adjectifs *noir* et *pollué* les locuteurs ont recours à une sorte de métaphore, mais cependant ils forment des collocations compositionnelles, par exemple, *des choses qui étaient noires* et *une langue polluée*. Il est aussi intéressant d'observer que l'adjectif *bête* est associé au nom inanimé *calcul* (dans *Mais on nous demande tellement de travail et je trouve qu'il y a beaucoup de travail qui/ qui pourrait ne pas être fait / Comme par exemple passer deux heures sur des/ des feuilles Excel pour faire des calculs bêtes*), bien que le dictionnaire (notamment Antidote) le considère comme un adjectif descriptif qui le plus souvent sert à exprimer la qualité d'un être humain.

Trois des collocations extraites peuvent être retrouvées dans le dictionnaire Antidote : *une très mauvaise opinion*, *un mauvais côté*, *des conditions épouvantables*. Les cinq collocations qui ne sont pas inclues dans le dictionnaire (*un français qui est mauvais*, *des choses qui étaient noires*, *des calculs bêtes*, *un français pollué*, *une langue polluée*) sont marquées, semble-t-il, par le désir des locuteurs de rendre leurs énoncés plus figurés afin d'exprimer leurs idées d'une manière plus claire.

4.2.3. Adjectifs dans les collocations illustrant la FL Magn

Une plus grande diversité d'adjectifs peut être notée dans les collocations illustrant la FL Magn, car dans les 57 collocations extraites on trouve 19 adjectifs différents : *acharné*, *cher*, *énorme*, *étendu*, *extrême*, *ferme*, *fort*, *global*, *grand*, *gros*, *important*, *long*, *majeur*, *principal*, *profond*, *puissant*, *riche*, *saillant*, *spécial*.

Dans la plupart des cas, les adjectifs relevés servent à exprimer la dimension des objets représentés par des noms concrets ainsi que l'intensité des phénomènes représentés par des noms abstraits, comme, par exemple dans *les grandes villes*, *des gros sacs à dos*,

phrases longues, *une énorme diversité*, *son opinion qui était très ferme*, *une amitié qui est profonde*, *un grand succès*, etc.

En outre, j'ai remarqué que les locuteurs choisissent l'adjectif *gros* non seulement pour parler des dimensions physiques d'un objet, mais aussi pour mettre en avant une ou des qualités caractéristiques de cet objet ou d'une personne. Notamment, en formant la collocation *un gros fumeur* le locuteur met l'accent sur le fait que quelqu'un fume beaucoup.

Le choix de présenter certaines collocations (telles que *les événements majeurs*, *spectacle principal*, *un rôle important*, *la langue principale*, etc.) comme celles illustrant la FL Magn peut paraître discutable. Cependant, les adjectifs utilisés dans les collocations citées ci-dessus semblent vraiment jouer le rôle d'intensificateurs. De plus, les définitions données aux FL concrètes au sein de la TST sont assez vastes, ce qui permet de considérer ces collocations comme les représentations de la FL Magn.

Il est à noter que la plupart des adjectifs extraits (39/50) peuvent être retrouvés dans le dictionnaire Antidote. Les collocations qui ne sont pas indiquées dans le dictionnaire sont principalement celles dont la base est un nom abstrait (par exemple*, son opinion qui était très ferme*, *une culture qui est assez riche*, *un phénomène quand même assez spécial*, *une énorme diversité*, *une très grande culture*).

4.2.4. Adjectifs dans les collocations illustrant la FL AntiMagn

En revanche, la diversité des adjectifs choisis par les locuteurs pour former des collocations représentant la FL AntiMagn est minimale. On ne trouve que 4 différents adjectifs parmi les 40 collocations extraites : *bas*, *léger*, *modeste*, *petit***.**

On peut noter que l'adjectif *petit* est utilisé non seulement pour parler de la dimension des objets (par exemple, dans *petites salles de concert*, *des petits podcasts*, *un petit exposé*), mais aussi pour former des diminutifs comme dans *petite langue*. Quant à l'adjectif *léger*, il fait partie de la collocation *un léger accent* dans laquelle il n'est pas employé dans le sens ʿ opposé à *lourd* ʾ, mais plutôt dans le sens ʿ peu perceptible ʾ, ʿ faible ʾ. C'est aussi le cas de l'adjectif *bas* utilisé dans la collocation *la basse classe* : il n'est pris dans le sens ʿ peu élevé ʾ, mais dans le sens ʿ qui est moindre en rang social ʾ.

19 collocations ne peuvent pas être retrouvées dans le dictionnaire Antidote. J'ai noté qu'il y avait une sorte d'incohérence. Par exemple, ce dictionnaire indique que le nom *accent* se combine avec l'adjectif *léger*, mais pas avec l'adjectif *petit*. De plus, on y retrouve des cooccurrences *un petit village*, *une petite île*, mais, en revanche, les collocations *les petites régions* et *ce petit territoire* ne sont pas inclues dans le dictionnaire, bien que les noms utilisés pour les former semblent se rapporter au même groupe thématique que *village* et *île*.

4.3. La fréquence des adjectifs

Ma recherche cherchait aussi à relever la fréquence des adjectifs utilisés pour former les collocations à l'étude. Les adjectifs les plus fréquents choisis par les locuteurs pour former des collocations illustrant les FL Bon, AntiBon, Magn et AntiMagn sont *bon* (15/50), *mauvais* (3/8), *grand* (20/57) et *petit* (36/40) respectivement. Le dictionnaire Antidote indique que la fréquence de *bon*, *grand* et *petit* est très haute et que l'adjectif *mauvais* est également assez fréquent. Ce sont des adjectifs très neutres et polyvalents dont le sens est le plus général. Le choix courant de ces quatre adjectifs s'explique, semble-t-il, par le style du discours présenté dans le corpus étudié. Il s'agit d'un registre

neutre, familier et informel qui présuppose l'utilisation des mots fréquents dans la langue.

La fréquence des autres adjectifs relevés est présentée dans le tableau ci-dessous :

Bon	**AntiBon**	**Magn**	**AntiMagn**
beau (11/50)	**pollué** (2/8)	gros (7/57)	bas (2/40)
apprécié (3/50)	bête (1/8)	**majeur** (4/57)	**léger** (1/40)
intéressant (3/50)	**épouvantable** (1/8)	**principal** (4/57)	**modeste** (1/40)
unique (3/50)	noir (1/8)	**important** (3/57)	
extraordinaire (2/50)		énorme (2/57)	
joli (2/50)		fort (2/57)	
sympa (2/50)		**global** (2/57)	
adorable (1/50)		long (2/57)	
blanc (1/50)		riche (2/57)	
enrichissant (1/50)		**acharné** (1/57)	
génial (1/50)		cher (1/57)	
grand (1/50)		**étendu** (1/57)	
haut (1/50)		**extrême** (1/57)	
magnifique (1/50)		**ferme** (1/57)	
positif (1/50)		profond (1/57)	
prisé (1/50)		**puissant** (1/57)	
		saillant (1/57)	
		spécial (1/57)	

Tableau 5. Fréquence des adjectifs utilisés dans les collocations extraites.

Il est à noter que 26 adjectifs relevés (à savoir ceux en caractères gras dans le Tableau 6) ne peuvent pas être retrouvés dans la liste de 907 mots de très haute fréquence qui a été proposée par Niklas-Salminen (cité par Çorbacioğlu Gönezer, 2015). Cette liste-là est basée sur un corpus oral enregistré dans les situations quotidiennes fréquentes au sein de la société française. Cependant, quelques thèmes proposés aux locuteurs lors de l'enregistrement du corpus FrancoToile sont parfois loin d'être quotidiens (la culture francophone, la langue française, etc.). Dans une certaine mesure, cela explique pourquoi les locuteurs ont de temps en temps recours aux adjectifs qui font penser au discours des personnes ayant un niveau d'éducation assez élevé (tels que *prisé*, *acharné*, *étendu*, *saillant*). Ainsi, les locuteurs ont recours à un éventail d'adjectifs assez grand même dans leurs énoncés oraux spontanés.

4.4. Choix des adjectifs selon les variables propres aux locuteurs

Finalement, cette recherche visait à mettre en lumière des différences quant au choix des adjectifs pour former ces collocations selon le sexe, l'âge, l'origine et le lieu de résidence des locuteurs. Plus bas, je présente des résultats pour chacune de ces variables séparément.

4.4.1. Sexe des locuteurs

Le nombre de femmes et d'hommes dont les énoncés sont présentés dans le corpus FrancoToile sont presque égaux (26 et 23 respectivement). De plus, le nombre de collocations extraites qui sont produites par les femmes et par les hommes sont aussi comparables : j'ai relevé 79 collocations formées par les femmes et 76 collocations formées par les hommes. Cela permet d'essayer de tracer des tendances quant au choix

des adjectifs par les locuteurs en fonction de leur sexe. J'ai pu faire deux observations à cet égard.

Premièrement, j'ai noté que les hommes dont les énoncés ont été enregistrés dans le cadre du projet FrancoToile ont recours à un éventail plus grand d'adjectifs que les femmes. Ainsi, les hommes ont utilisé 36 adjectifs des 52 adjectifs qui apparaissent dans les collocations extraites, tandis que les femmes n'en ont employé que 28. Comme le niveau d'éducation, la profession et d'autres critères du statut social ne sont pas pris en considération dans le cadre de cette recherche et que l'envergure du corpus étudié n'est pas assez grande, il est impossible d'aboutir à une conclusion à ce propos.

Deuxièmement, il est intéressant de constater que les femmes ont recours aux adjectifs les plus fréquents presque aussi couramment que les hommes. La répartition des adjectifs les plus utilisés selon le sexe des locuteurs est présentée dans le tableau suivant[13] :

	Femmes	**Hommes**
Petit	21/79	15/76
Grand	8/79	12/76
Bon	8/79	7/76

Tableau 6. Répartition des trois adjectifs selon le sexe des locuteurs.

Dans cet ordre d'idées, on peut noter que dans leurs énoncés oraux spontanés les locuteurs des deux sexes ont tendance à choisir des adjectifs qui appartiennent au vocabulaire de base et qui sont considérés par les dictionnaires comme des adjectifs de grande fréquence. C'est, peut-être, dû au fait qu'il s'agit de discours spontané qui n'était

[13] L'adjectif *mauvais* n'est pas pris en compte, car on ne le retrouve que dans les énoncés de trois locuteurs.

pas préparé. De plus, cela semble aussi s'expliquer par le fait que l'enregistrement des énoncés se passait en présence de l'interviewer ce qui incitait les locuteurs de choisir les mots neutres plutôt que les mots expressifs.

4.4.2. Âge des locuteurs

Quant à l'âge des locuteurs, ce n'est pas facile de relever des tendances, car le nombre de locuteurs appartenant aux différentes tranches d'âge n'est pas vraiment comparable. Toutefois, on peut remarquer que la répartition des adjectifs courants, tels que *petit*, *beau*, *bon*, *grand* et *gros* est presque la même pour chaque groupe. De plus, j'ai noté que les jeunes enfants (de 0 à 9 ans) ne sont pas vraiment enclins à utiliser des collocations dans leurs énoncés, ce qui n'est pas le cas des personnes âgées qui ont 65 ans et plus. Ainsi, dans les énoncés de trois enfants qui ont 9 ans et moins, je n'ai relevé qu'une seule collocation, tandis que dans les énoncés de trois locuteurs de 65 ans plus, il y avait 18 collocations formées avec 9 adjectifs différents. Il est à noter qu'un des locuteurs qui a plus de 65 ans est auteur et éditeur. Il est donc fort probable que cet homme s'exprime d'une manière plus sophistiquée que les autres et que son discours est plus varié que celui des autres locuteurs. De plus, les personnes âgées, ayant le niveau d'éducation plus élevé que celui des jeunes enfants, ont, semble-t-il, un vocabulaire plus large. J'ai aussi observé que la diversité des adjectifs utilisés tend à augmenter proportionnellement avec l'âge des locuteurs. Ainsi, les jeunes locuteurs qui ont entre 19 et 25 ans ont formé 68 collocations à l'aide de 27 adjectifs différents (soit moins de la moitié), les locuteurs qui ont entre 26 et 39 ans ont formé 14 collocations en ayant recours à 6 différents adjectifs (presque la moitié), les locuteurs qui ont entre 55 et 64 ans ont formé 21 collocations avec 13 différents adjectifs (plus que moitié), les locuteurs qui

ont entre 55 et 64 ans ont formé 26 adjectifs en utilisant 18 différents adjectifs (plus que la moitié), les locuteurs qui ont 65 ans et plus ont formé 18 collocations au moyen de 9 différents adjectifs (moitié). C'est, semble-t-il, très naturel que la manière de s'exprimer dépende beaucoup de l'âge du locuteur, ce qui implique aussi son éducation, l'expérience vécue et de nombreuses variables qui ne sont pas pris en compte au sein de ce projet.

4.4.3. Origine des locuteurs

Bien que dans le corpus étudié il y ait des énoncés des locuteurs qui proviennent de plusieurs pays, ce sont seulement le nombre de locuteurs d'origine française (24/49) et le nombre de locuteurs d'origine canadienne (20/49) qui sont comparables. J'ai constaté que les locuteurs originaires de France et les locuteurs d'origine canadienne ont formé un nombre de collocations presque égal : 67 et 65 respectivement. La diversité des adjectifs utilisés est aussi presque similaire chez les locuteurs de ces deux groupes : les locuteurs d'origine française ont recours à 25 adjectifs différents et ceux d'origine canadienne forment des collocations au moyen de 22 adjectifs différents. Cependant il y a 14 adjectifs qui ne sont utilisés que par les locuteurs d'origine française (*génial*, *joli*, *magnifique*, *prisé*, *sympa*, *cher*, *ferme*, *profond*, *riche*, *bête*, *épouvantable*, *bas*, *modeste*, *acharné*) et 12 adjectifs utilisés seulement par les locuteurs d'origine canadienne (*unique*, *pollué*, *léger*, *enrichissant*, *extraordinaire*, *global*, *long*, *majeur*, *principal*, *puissant*, *saillant*, *spécial*). J'ai noté qu'il y a un seul cas où les locuteurs d'origine française et d'origine canadienne associent des adjectifs différents au mêmes nom en formant les collocations presque synonymiques. Chez un locuteur provenant de France on trouve la collocation *une culture qui est assez riche*, tandis qu'un locuteur d'origine canadienne a formé la collocation *une très grande culture*.

On peut également faire quelques remarques en ce qui concerne les adjectifs de grande fréquence. Premièrement, chez des locuteurs d'origine française l'adjectif *gros* est associé à plusieurs noms différents (*un gros fumeur*, *des gros trous*, *un gros mix*, *les deux grosses pôles*[14], *des gros sacs à dos*) et exprime non seulement la dimension physique d'un objet, mais joue aussi le rôle d'intensificateur. Par contre, un locuteur d'origine canadienne qui utilise cet adjectif le choisit pour souligner l'importance de l'Acadie (*l'Acadie c'est une grosse chose*). Deuxièmement, les locuteurs provenant des deux pays utilisent l'adjectif *important* presque de la même manière en lui attribuant le sens ʻ qui a des conséquences considérables ʼ (*des sessions plus importantes*, *un message qui est important*, *un rôle important*). Troisièmement, un locuteur d'origine française utilise l'adjectif *fort* dans le sens ʻ difficile à détruire ʼ (*des liens forts*), tandis qu'un locuteur d'origine canadienne l'emploie comme un synonyme de *remarquable* (*un accent fort*). Quatrièmement, l'adjectif *grand* est associé au nom *succès* dans le discours des locuteurs provenant des deux pays.

Ainsi, il n'y a pas de différence globale quant au choix des adjectifs par les locuteurs provenant des pays différents. Cela est probablement dû à l'envergure du corpus étudié qui contient les énoncés d'un nombre assez limité de locuteurs. En outre, dans le corpus étudié, on ne retrouve pas de régionalismes auxquels on pourrait s'attendre vu la diversité des origines des locuteurs. On peut supposer que les locuteurs, conscients du fait que leurs propos étaient enregistrés et de l'objectif principal du corpus (présenter les opinions des francophones du monde entier sur la francophonie), se sont exprimé dans

[14] La forme du féminin est choisie par le locuteur.

un français plutôt standard qui peut être compris par tous les locuteurs de cette langue, peu importe leur origine.

J'ai pu aussi relever une particularité propre aux locuteurs d'origines différentes. J'ai noté que le locuteur d'origine suisse essaye de rendre son discours plus métaphorique et forme les collocations *des choses qui étaient blanches* et *des choses qui étaient noires* en parlant de l'éducation d'avant. On retrouve aussi cette tendance à parler d'une manière plus métaphorique chez un locuteur d'origine acadienne qui utilise l'adjectif *pollué* dans les collocations *un français pollué* et *une langue polluée*. Bien sûr, l'exemple d'un seul locuteur ne me permet pas de généraliser sur la manière métaphorique de parler chez les locuteurs d'origine suisse ou acadienne.

4.4.4. Lieu de résidence des locuteurs

Dans le corpus FrancoToile, il y a des énoncés des locuteurs qui vivent dans un environnement principalement anglophone : d'une locutrice résidant en Californie et de huit personnes résidant à Victoria. On peut donc supposer que la langue anglaise produit un impact sur la manière dont ces locuteurs s'expriment et choisissent des adjectifs pour former les collocations à l'étude. J'ai constaté que le discours de la locutrice qui habite en Californie n'est pas marqué par l'influence de l'anglais. En revanche, les collocations *une vision globale* et *les événements majeurs* formées par deux locuteurs résidant à Victoria peuvent être vues comme des calques des syntagmes lexicaux anglais *global vision* et *major events*. Cependant, dans le dictionnaire Antidote, les collocations *vision globale* et *événement majeur* sont indiquées parmi les cooccurrences les plus fréquentes pour les noms *vision* et *événements*. Cela pourrait signifier que ces deux syntagmes lexicaux sont

bien ancrés dans la langue française et ne font presque pas penser aux influences de l'environnement anglophone.

Le fait qu'on ne trouve pas de dissemblances majeures quant au choix des adjectifs par les locuteurs provenant des pays différents et résidant dans les régions différentes s'explique, semble-t-il, par le fait que les locuteurs étaient conscients que leurs propos étaient enregistrés et qu'ils essayaient, peut-être, de cacher leurs particularités régionales et de s'exprimer aussi correctement que possible. De plus, le corpus étudié ne représente le discours que d'un nombre limité de locuteurs.

5. Conclusion

La présente recherche, inscrite dans le cadre de la TST, cherchait à mettre en lumière les adjectifs que les locuteurs du français emploient dans leurs énoncés oraux en formant des collocations qui illustrent les FL Bon, AntiBon, Magn et AntiMagn. Au cours de mon étude, j'ai tenté de répondre aux questions suivantes : 1. Quels adjectifs les locuteurs du français utilisent-ils dans leurs énoncés oraux pour former des collocations représentant les FL Bon, AntiBon, Magn et AntiMagn ? 2. Quelle est la fréquence des adjectifs utilisés pour former les collocations en question ? 3. Observe-t-on des différences quant au choix des adjectifs pour former ces collocations selon l'origine géographique des locuteurs, ainsi que selon d'autres variables, comme, par exemple, l'âge des locuteurs ?

Comme source de données à analyser j'ai choisi le corpus FrancoToile qui recueillit les énoncés de 49 locuteurs provenant de plusieurs pays et résidant dans les régions différentes du monde. Le choix du corpus a imposé quelques limites à la recherche. Premièrement, bien que les énoncés présentés dans le corpus ne soient pas préparés, les locuteurs étaient conscients que leurs propos étaient enregistrés. Deuxièmement, les thèmes de la conversation n'étaient pas choisis librement par les locuteurs, mais c'était l'intervieweur qui leur proposait de s'exprimer sur tel ou tel sujet. Troisièmement, le nombre de locuteurs appartenant aux groups différents selon les variables (sexe, âge, origine, lieu de résidence) n'était pas toujours comparable, ce qui ne m'a pas permis d'aboutir à des conclusions généralisantes et de relever des tendances. Quatrièmement, le corpus ne donne pas d'information sur le niveau d'éducation des locuteurs qui peut avoir une influence notable sur leur façon de s'exprimer. Cependant,

malgré toutes ces limitations, j'ai pu faire quelques remarques concernant des adjectifs utilisés dans les collocations qui représentent les FL Bon, AntiBon, Magn et AntiMagn.

J'ai extrait 155 collocations du type « adjectif + nom » et « nom + adjectif » dont 50 collocations illustrent la FL Bon, 8 collocations illustrent la FL AntiBon, 57 collocations représentent la FL Magn et 40 collocations représentent la FL AntiMagn. En formant les collocations illustrant la FL Magn les locuteurs ont eu recours au plus grand éventail d'adjectifs (19 différents adjectifs sur 57 collocations), tandis que la diversité des adjectifs dans les collocations représentant la FL AntMagn est minimale (4 différents adjectifs sur 40 collocations). Dans l'ensemble, on peut dire que les locuteurs font parfois recours aux syntagmes lexicaux qui ne sont pas retenus dans le dictionnaire : 67 des collocations extraites ne peuvent pas être retrouvées dans le dictionnaire Antidote.

Quant à la fréquence des adjectifs, les locuteurs utilisent le plus souvent des mots polyvalents dont le sens est très large (à savoir *bon*, *mauvais*, *grand* et *petit*) et dont l'utilisation est tout à fait naturelle quand il s'agit d'un registre plutôt neutre et informel. Cependant, les locuteurs ont parfois recours aux adjectifs qui ne font pas partie des mots les plus fréquents da la langue française (par exemple, *prisé*, *acharné*, *étendu*, *saillant*) ce qui s'explique, peut-être, par le niveau d'éducation des locuteurs ou par le thème de la conversation proposé.

En ce qui concerne les différences quant aux choix des adjectifs selon les variables propres aux locuteurs, il n'y en a pas beaucoup. En comparant les collocations formées par les hommes et les femmes, j'ai noté que dans le discours spontané les locuteurs des deux sexes ont tendance à choisir les adjectifs qui sont considérés par les dictionnaires comme des mots de haute fréquence. J'ai observé la même situation en

comparant les collocations formées par les locuteurs appartenant aux différentes tranches d'âge : la répartition des adjectifs courants, tels que petit, beau, bon, grand et gros est presque identique pour chaque groupe. De plus, j'ai remarqué que plus les locuteurs étaient âgés, plus ils avaient tendance à utiliser des collocations dans leurs énoncés et plus la palette des adjectifs employés était grande.

Bien que le corpus étudié contienne des énoncés des locuteurs provenant de régions différentes de la francophonie, ce sont seulement le nombre d'originaires de France et le nombre de locuteurs d'origine canadienne qui sont comparables. Certains adjectifs n'étaient utilisés que par les locuteurs appartenant à un groupe, tandis que d'autres (notamment les adjectifs de haute fréquence) peuvent être retrouvés dans les énoncés d'originaires des deux pays. Pourtant, j'ai pu remarquer quelques différences mineures quant à la formation des collocations : les bases des collocations formées par les locuteurs d'origine française et par ceux d'origine canadienne ne sont pas toujours les mêmes, ce qui a un impact sur le sens du collocatif (cf. *des liens forts* et *un accent fort*). Quant au lieu de résidence, je n'ai pas observé d'influence nette de l'environnement linguistique à laquelle on pourrait s'attendre dans les énoncés des locuteurs résident dans les régions principalement anglophones. Cela est probablement dû aux limitations imposées par le choix du corpus qui sont décrites ci-dessus.

Dans l'ensemble, cette recherche contribue à la compréhension du fonctionnement du lexique au sein du français parlé. De plus, j'ai eu recours aux FL qui ne sont pas souvent utilisées dans le cadre de la linguistique de corpus. Toutefois, cet instrument semble être assez efficace pour la classification du vocabulaire abstrait et descriptif. Ainsi, les FL pourraient être utilisées dans l'enseignement du français afin de

faire connaître aux apprenants des collocations possibles dans cette langue. Une étude pareille basée sur un corpus de plus grande envergure sera nécessaire à cette fin.

Bibliographie

Alonso Ramos, M. (2000). Critères heuristiques pour l'encodage des collocations au moyen de fonctions lexicales. *EURALEX, 9*, 463-473.

Antidote Mobile – Français (version 3.2.5) [logiciel]. Récupéré de https://apps.apple.com/app/id797819487

Benigno, V., Grossmann, F. et Kraif, O. (2015). Les collocations fondamentales : une piste pour l'apprentissage lexical. *Revue française de linguistique appliquée, 20*(1), 81-96.

Bolly, C. (2010). Flou phraséologique, quasi-grammaticalisation et pseudo marqueurs de discours : un no man's land entre syntaxe et discours ? *Linx, 62*, 11-38. Récupéré de https://journals.openedition.org/linx/1356

Çorbacioğlu Gönezer, E. (2015). Le français fondamental et la statistique lexicale et l'utilisation des mots fréquents dans les manuels préparés pour FLE. *Journal of international social research*, *8*(41), 941-948. Récupéré de http://www.sosyalarastirmalar.com/cilt8/sayi41_pdf/5egitim/corbacioglugonezer_eda.pdf

Corpus FrancoToile. (2008-2014). Accessible sur le site du projet : https://francotoile.uvic.ca/index.xql

Diachenko, P. (2006). Lexical functions in learning the lexicon. Dans A. Méndez-Vilas (dir.), *Current developments in technology-assisted education* (p. 538-542). Badajoz, Espagne Formatex.

Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. Dans F. Palmer (dir.), *Selected papers of J. R. Firth 1952–1959* (p. 168-205). Londres, R.-U. : Longman.

Gries, S. (2013). 50-something years of work on collocations. What is or should be next… *International Journal of Corpus Linguistics, 18,* 137-165. Récupéré de https://www.academia.edu/23085867/50-something_years_of_work_on_collocations_What_is_or_should_be_next_

Kahane, S. (2003). The Meaning-Text theory. Dans V. Ágel, L. M. Eichinger, H. W. Eroms, P. Hellwig, H. Jürgen Heringer et H. Lobin (dir.), *Dependency and valency. An international handbook of contemporary research* (p. 546-570). Berlin, Allemagne : De Gruyter. Récupéré de https://hal-univ-paris10.archives-ouvertes.fr/hal-02293104/document

Legallois, D. et Tutin, A. (2013). Présentation : vers une extension du domaine de la phraséologie. *Langages, 189*(1), 3-25. https://doi.org/10.3917/lang.189.0003

L'Homme, M.-C. (2002). Fonctions lexicales pour représenter les relations sémantiques entre termes. *Traitement automatique de la langue, 43*, 19-41.

Mel'čuk, I. (1974). Esquisse d'un modèle linguistique du type "Sens⇔Texte". Dans F. Bresson et J. Mehler (dir.), *Problèmes actuels en psycholinguistique* (p. 291-317). Paris, France : Éditions du Centre national de la recherche scientifique.

Mel'čuk, I. (1981). Meaning-Text models: A recent trend in soviet linguistics. *Annual Review of Anthropology, 10*, 27-62.

Mel'čuk, I. (1982). Lexical functions in lexicographic description. Dans *Proceedings of the Eights annual meeting of the Berkeley linguistics society* (p. 427-444).

Berkeley, CA: University of California.

Mel'čuk, I. (1985). Semantic Bases for Linguistic Description (Meaning-Text Theory). Dans *Proceedings of the Twelfth LACUS Forum* (p. 41-87).

Mel'čuk, I. (1988). *Dependency syntax: Theory and practice*. New York, NY: State University of New York Press.

Mel'čuk, I. (1993). *Cours de morphologie générale. Vol. 1 : Introduction et première partie : le mot*. Montréal, Canada : Les Presses de l'Université de Montréal.

Mel'čuk, I. (1997) Vers une linguistique Sens-Texte. Leçon inaugurale (donnée le 10 janvier 1997), Collège de France, Chaire internationale. Récupéré de http://olst.ling.umontreal.ca/pdf/melcukColldeFr.pdf

Mel'čuk, I. (1998). Collocations and lexical functions. Dans A.P. Cowie (dir.), *Phraseology. theory, analysis, and applications* (p. 23-53). Oxford : Clarendon Press.

Mel'çuk, I. (1999). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques.* Montréal, Canada: Les Presses de l'Université de Montréal.

Mel'čuk, I. (2003). Collocations dans le dictionnaire. Dans T. Szende (dir.), *Les écarts culturels dans les dictionnaires bilingues* (p. 19-64). Paris, France : Honoré Champion.

Mel'čuk, I. (2012). Phraseology in the language, in the dictionary, and in the computer. Dans H. Buhofer (dir.), *Yearbook of phraseology* (p. 31-56). Berlin, Allemagne : Mouton de Gruyter.

Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur phrasèmes, mais... *Cahiers de*

lexicologie, 102(1), 129-149. Récupéré de http://olst.ling.umontreal.ca/pdf/MelcukPhrasemes2011.pdf

Mel'čuk, I., Clas, A. et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Gembloux, Belgique : Éditions Duculot.

NVivo (version 12) [logiciel]. Récupéré de https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/support-services/nvivo-downloads

Polguère, A. (1998) La théorie Sens-Texte. *Dialangue, 8-9*, 9-30. Récupéré de http://olst.ling.umontreal.ca/pdf/PolgIntroTST.pdf

Polguère, A. (2003a). Collocations et fonctions lexicales : pour un modèle d'apprentissage. Dans F. Grossmann et A. Tutin (dir.), *Les collocations : analyse et traitement* (p. 117-133). Amsterdam, Pays-Bas : De Werelt.

Polguère, A. (2003b). *Lexicologie et sémantique lexicale : notions fondamentales.* Montréal, Canada : Presses de l'Université de Montréal.

Polguère, A. (2011). Perspective épistémologique sur l'approche linguistique Sens-Texte. *Mémoires de la Société de Linguistique de Paris*, *20*, 79-114.

Tutin, A. et Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue française de linguistique appliquée, 7*(1), 7-25.

Annexe 1. Toutes les collocations extraites en contexte selon les FL.[15]

FL Bon

chez nous **les fêtes folkloriques sont très appréciées** (France/F/55-64)

chevaux qui sont harnachés avec **de très beaux harnais** très décorés (France/F/55-64)

nous avons également **de très bons chanteurs** en langue provençale **qui sont aussi très prisés** par nous bien-sûr (France/F/55-64)

notre petite **langue chantante est appréciée** (France/F/55-64)

l'anglais parce que c'est **une matière qui est très intéressante** (France/H/10-14)

depuis deux ans les expériences que l'on fait sont très bien appréciées et acceptées (Canada/H/55-64)

alors on eu **des très bonnes discussions** parce que il y avait d'autres pays francophones (Canada/H/55-64)

Alors c'était une expérience très enrichissante (Canada/H/55-64)

ce que j'aime bien ici c'est que les gens sont très sympas (France/H/19-25)

elles ont toujours **des jolis sacs** (France/H/19-25)

en France on a **de très bons professeurs** mais ils prennent beaucoup moins soin de nous (France/H/19-25)

mais **la vie était plus belle** on n'avait pas de sous mais c'était plus beau que maintenant (France/F/65+)

je mange à la cantine tous les jours […] **les dames de la cantine sont sympas** (France/F/10-14)

il y a des chansons mais qui sont géniales quoi (Canada/F/19-25)

ce sont **des bons chanteurs** quoi donc je lui ai fais partager ça (Canada/F/19-25)

les gens qui te diront oui le paysage est magnifique, les gens sont adorables (Canada/F/19-25)

mais ils ont une très belle image en général de la France (Canada/F/19-25)

il y avait **des choses qui étaient blanches** il y avait des choses qui étaient noires (France/H/40-54)

[15] Les collocations qui ne peuvent pas être retrouvées dans le dictionnaire Antidote sont en caractères gras. Les collocations sont suivies par un code indiquant le lieu de résidence, le sexe et la tranche d'âge des locuteurs qui ont formés ces syntagmes.

l'aspect positif … le Français est perfectionniste […]donc oui il est en général doué dans certains domaines (France/H/40-54)

il a **une belle capacité** d'adaptation je crois en général (France/H/40-54)

on a reçu une bonne éducation ça a bien changé maintenant (France/H/40-54)

on a reçu une éducation qui était quand même d'un bon niveau même si j'ai pas fait de **très hautes études** (France/H/40-54)

Et en fait **cet accent marseillais qui était très joli** à la base est en train de changer (France/H/19-25)

je crois que je ne m'étais jamais interrogée sur ce qu'était la francophonie (France/F/26-39)

je trouve c'est genre comme un art, **chaque art est unique** donc **chaque langue est unique** aussi (Canada/F/19-25)

une autre chose aussi qui est intéressante c'est…on a des cours d'initiation…dès le début (Canada/F/19-25)

je dirais que c'est *Le Jardin de mon père*, c'est **un beau camping** (Canada/F/10-14)

je trouve que c'est un c'est un outil extraordinaire pour pour les francophones et francophiles (Canada/F/40-54)

qui est une de nos grandes chanteuses dont on est très très fiers, **qui a une voix extraordinaire** (Canada/F/40-54)

un artiste québécois […] Samian […] qui a vraiment des beaux textes (Canada/F/40-54)

nos interprétations de la musique sont des interprétations […] très uniques (Canada/H/55-64)

je pense ça une belle langue (Canada/H/19-25)

on va plus dire des belles phrases longues et romantiques ou on va plus dire des belles phrases explicatives (Canada/F/19-25)

je trouve que le français c'est une belle langue (Canada/F/19-25)

je trouve que tous les pays ont des belles langues (Canada/F/19-25)

mais est-ce qu'il y a un bon côté à l'évolution (Canada/F/19-25)

cette année-là c'était vraiment la meilleure année (Canada/H/19-25)

ils passent **du bon temps** ensemble (Canada/F/26-39)

en échange ils leur donnaient du bon manger (Canada/H/65+)

donc les meilleurs classements ont **les meilleures écoles** secondaires (Canada/F/19-25)
un examen vous permet d'accéder à **une meilleure école** ou à **des meilleurs professeurs** (Canada/F/19-25)
les encourager au niveau pédagogique de donner **des meilleurs cours** (France/H/19-25)

FL AntiBon

on utilise un français que certains disent un peu **un français pollué**, **un français qui est mauvais** (Canada/H/19-25)
Je dirais pas que c'est **une langue polluée** (Canada/H/19-25)
il y avait des choses qui étaient blanches, il y avait **des choses qui étaient noires** (France/H/40-54)
passer deux heures sur des feuilles Excel pour faire des calculs bêtes (France/H/19-25)
ils avaient une très mauvaise opinion d'eux-mêmes (France/H/65+)
est-ce qu'il y a un mauvais côté(Canada/F/19-25)
il restait dans un camp external link dans le bois dans des conditions épouvantables (Canada/H/65+)

FL Magn

c'est **un message qui est important** pour les parents (Canada/H/55-64)
on a **un phénomène quand même assez spécial** (Canada/H/55-64)
C'est en s'impliquant dans les événements majeurs anglophones (Canada/H/55-64)
à l'intérieur de **ces activités majeures**, de ces événements majeurs (Canada/H/55-64)
essayer de s'impliquer dans tous ces événements majeurs-là (Canada/H/55-64)
Radio-Radio qui était ici en **spectacle principal** (Canada/H/55-64)
ici **les cafés sont tous grands** (France/H/19-25)
elles mettent des Ugg, des leggings, des gros sacs à dos (France/H/19-25)
il avait son **opinion qui était très ferme** (Canada/F/19-25)
tu crées très vite des liens et des liens forts (Canada/F/19-25)
tout à un prix hein et le prix de la Provence est très cher (France/F/55-64)
acquérir une amitié qui est profonde (France/F/55-64)
il y a de grandes valeurs qui demandent à être respectées (France/F/55-64)

dealer avec **le problème de classement et de compétition qui est assez extrême** (Canada/F/19-25)

je pense qu'on a **une culture qui est assez riche** (France/F/19-25)

c'est **une culture qui est assez riche** (France/F/19-25)

à travers les grands rois de France (France/F/19-25)

à travers tout ce qui est les grandes chaînes (France/F/19-25)

on a toujours un moment assez long pour manger (France/H/10-14)

certaines salles musicales dans les grandes villes (France/F/26-39)

d'avoir une très très grande diversité linguistique (Canada/H/19-25)

la francophonie maintenant c'est **une énorme diversité** (Canada/H/19-25)

Je crois que ça a une influence énorme sur les cultures (Canada/H/26-39)

il y a pas vraiment une grande différence au niveau au niveau des goûts (Canada/H/26-39)

Qui a […] une voix aussi puissante sinon plus que Céline Dion (Canada/F/40-54)

la radio communautaire joue vraiment un rôle important pour la motivation (Canada/F/40-54)

c'est un fait saillant de la musique traditionnelle (Canada/H/55-64)

qui a une vision globale (Canada/H/55-64)

j'ai une vision globale en ce qui concerne la musique aussi (Canada/H/55-64)

c'était un grand succès (Canada/H/55-64)

un an d'expérience au sein d'une grande entreprise mondiale (France/H/19-25)

c'est déjà une très grande entreprise pour la France (France/H/19-25)

il y a **une grande récréation** et on sort à seize heures trente (France/F/10-14)

une grande tradition c'est la sieste comme en Espagne (France/H/26-39)

une très grande culture au niveau de la musique (Canada/H/26-39)

on a une très grande variété aussi de musique (Canada/H/26-39)

l'île d'Yeu s'y est reconnue et a ménagé un très grand succès à cette petite entreprise (France/H/65+)

c'était un grand grand succès (France/H/65+)

À cette époque je suis encore un gros fumeur (France/H/65+)

c'est deux années de travail acharné généraliste (France/H/19-25)

avec de temps en temps **des sessions plus importantes** de ce qu'on appelle les devoirs surveillés (France/H/19-25)

le français est la langue principale parlée (Canada/F/19-25)

n va plus dire des belles phrases longues (Canada/F/19-25)

la langue principale parlée est le français (Canada/F/19-25)

prioriser une langue qui n'est même pas notre langue principale au Québec (Canada/F/19-25)

l'Acadie c'est **une grosse chose** (Canada/F/19-25)

mais on fait un grand effort dans la famille (Californie/F/40-54)

Alors pour nous c'est une grande importance (Californie/F/40-54)

et au final c'est **un gros mix** je pense (Canada/F/19-25)

c'est **les deux grosses pôles** on va dire (Canada/F/19-25)

je crois que c'est ça un gros problème d'éducation (France/H/40-54)

Donc c'est **un grand mélange** je crois, **un grand mélange** (France/H/40-54)

je l'imagine comme un territoire donc assez étendu finalement (France/F/26-39)

on faisait des gros trous (France/F/10-14)

les gens avaient un accent fort vraiment typique de la région (France/F/19-25)

FL AntiMagn

les écoles sont trop petites, refusent du monde (Canada/H/55-64)

Et pis te parler avec un léger accent comme ça (Canada/H/55-64)

on est allées dans les magasins acheter des petits cadeaux (Canada/F/0-9)

ils se disaient c'était la basse classe, plus bas que la basse classe (Canada/H/65+)

une femme qui avait deux valises et puis une petite fille (Canada/H/65+)

alors il prend la petite fille une valise et puis il passe à travers devant la police (Canada/H/65+)

Notre **petite langue** chantante est appréciée (France/F/55-64)

qui existe dans mon pays, l'île Maurice, une petite îledans l'océan Indien (Canada/F/19-25)

on peut dire que ch'suis **une petite voyageuse** vu que je me promène un petit peu partout (Canada/F/19-25)

j'espère que ça vous a donné une petite idée ou ça vous a fait réfléchir (Canada/F/19-25)

je vais commencer par vous expliquer une petite anecdote qui m'est arrive (Canada/F/19-25)

je trouve que **la petite communauté** tout le monde s'aide, tout le monde est bien ensemble (Canada/F/19-25)

les petites business c'est toute du monde comme local qui ont ça (Canada/F/19-25)

je viens **d'une petite région** dans le fin fond du Québec

ce sont souvent de **petites communautés** francophones (Canada/F/26-39)

Mais on a aussi **notre petite langue, petit dialecte** qu'on appelle (Canada/F/19-25)

je me souviens juste du petit livret qu'on avait et on chantait (Canada/F/19-25)

c'est **des petites expressions** qui restent et qui t'expliquent des faits de vie quoi (Canada/F/19-25)

cette expérience très modeste concerne un journal que j'ai eu l'occasion de réaliser (France/H/65+)

est-ce que l'un de vos gars, **une petite équipe** pourrait venir m'interviewer? (France/H/65+)

D'autant plus que sur **ce petit territoire** de cinq kilomètres de large sur neuf kilomètres de long (France/H/65+)

Ceci pour dire à quel point […] **un petit territoire** comme la France peut receler d'accents (France/H/65+)

Si donc sur **un si petit territoire** vous avez des inflexions (France/H/65+)

Qui sont finalement des petites salles de concert qui vont accueillir des artistes émergents (France/F/26-39)

Il y a beaucoup de festivals aussi dans même dans **les petites regions** (Canada/H/26-39)

je pense ma mère elle a conté un petit mensonge euh elle a dit quelque chose (Canada/H/19-25)

il est bien mignon **ton petit accen**t mais je comprends absolument rien (Canada/F/19-25)

Ils apportent tous à l'université leur petit sandwich, ils mangent en classe (France/H/19-25)

souvent on prenait un petit café, des petits cafés d'ailleurs parce que ici les cafés sont tous grands (France/H/19-25)

je viens d'un petit village du Québec, qui s'appelle Sainte-Apolline, comté de Montmagny (Canada/F/40-54)

de se donner un but euh de parler, de faire **un petit exposé** oral (Canada/F/40-54)

on fait **des petits podcasts** et on peut les faire jouer à la radio (Canada/F/40-54)

on avait quatre matches dans la journée mais **des petits matches** d'une demi-heure (Canada/H/10-14)

j'habite à Haute-Aboujagane, c'est un petit village près de Shédiac au Nouveau-Brunswick (Canada/F/19-25)

sans que ce soit organisé dans le cadre d'un petit voyage avec que des Français (France/F/26-39)