



Univerzita Karlova v Praze

Filozofická fakulta

Ústav Románských studií / Francouzská filologie

Bakalářská práce



Kristína Konštiaková

Différents emplois du mot *comme* en français contemporain

Construction « *être* + **ADJ** + *comme* »

Typy užití slova *comme* v současné francouzštině

Konstrukce « *être* + **ADJ** + *comme* »

The Use of the Word *comme* in contemporary French

Construction « *être* + **ADJ** + *comme* »

Praha 2016

Vedoucí práce: PhDr. Olga Nádvoříková, Ph.D.

Remerciements

Je voudrais adresser mes sincères remerciements à la directrice de mon mémoire de licence PhDr. Olga Nádvorníková, Ph.D. pour tout son temps dédié à ce travail, ses idées et ses précieux conseils qui l'ont guidé et enrichi. Je tiens à remercier aussi mon ancien professeur du lycée bilingue Mgr. Nicolas Theys de sa complaisance de relire et corriger mon manuscrit, ainsi que mon petit ami Dejan, mes amis et ma famille pour leurs encouragements.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 17. května 2016

.....

Kristína Konštiaková

Abstrakt

Bakalářská práce se zabývá francouzským slovem *comme*, které se vzhledem k jeho polysémantičnosti jeví problematické, co se týče zařazení jako slovního druhu. Cílem teoretické části je zmapovat všechny užití tohoto slova s využitím referenčních slovníků a odborné literatury. Praktická část se podrobněji věnuje užití *comme* ve srovnáních. Pomocí jazykových korpusů se zkoumá schéma, kde je *comme* v pozici za slovesem *být* a adjektivem (*être* + ADJ + *comme*). Toto schéma je zkoumané z frekvenčního hlediska v subkorpusech představujících různé typy textů. Součástí frekvenční analýzy je rovněž kvalitativní analýza pravého kontextu slova. Dalším předmětem zájmu je identifikace výskytů s přeneseným významem a vyhledání nejfrekventovanějších adjektiv jako kandidátů pro tvorbu kolokací, případně frazémů. Nakonec se potenciální kolokace porovnávají se slovníky, aby se potvrdila, respektive aby se našla nová zatím nelexikalizovaná spojení.

Klíčová slova: *comme*, srovnání, korpus, typ textu, kolokace

Abstract

The work is dealing with the French word *comme*, whose polysemy involves a problematic determination of the part of speech. The aim of theoretical part is to map all the cases of use of this word by using reference dictionaries and specialized literature. The practical part examines in more detail the use of *comme* in comparison. By means of language corpora, *comme* is studied in the construction where preceded by the verb *be* and an adjective (*être* + ADJ + *comme*). This structure is analyzed from the frequency point of view in subcorpora representing different types of texts. The frequency analysis includes qualitative analysis of the right context. Another goal is to identify the occurrences with a figurative meaning and to search for the most frequent adjectives as candidates figuring in collocations or perhaps even in idioms. Finally, the potential collocations are contrasted with dictionaries either to confirm them or to discover new ones, yet not lexicalized.

Key words: *comme*, corpus, text type, comparison, collocation

Mots-clés : *comme*, corpus, type de texte, comparaison, locution figée

Liste des abréviations

ADJ = adjectif

ccir. = complément circonstanciel

ipm = instances per million

occ. = occurrence

PP = participe présent/passé

prop. = proposition

SN = syntagme nominal

SV = syntagme verbal

Sommaire

Introduction	9
1 Comme et ses emplois	11
1.1 Emplois selon les dictionnaires	11
1.2 Emplois selon les grammaires.....	12
1.3 Difficultés du classement	13
1.4 Emplois secondaires.....	14
2 Comme conjonction de subordination	17
2.1 Comme conjonction de temps	17
2.2 Comme connecteur de cause	18
3 Comme adverbe de manière	21
3.1 Exclamation d'intensité.....	21
3.2 Comparaison	22
3.2.1 Types de comparaison	23
3.2.2 Comparaison avec <i>comme</i>	24
3.2.3 <i>Comme</i> + X	25
3.2.4 Comparaison métaphorique	27
3.2.5 Locutions figées.....	29
4 Construction « être + ADJ + comme »	31
4.1 Hypothèse et méthodologie.....	31
4.2 1 ^{ère} phase – « être + ADJ + comme + X ».....	33
4.2.1 Fréquence de la construction selon le type de texte.....	33
4.2.2 Fréquence de la construction selon le genre littéraire.....	38
4.2.3 Résultats « intrus ».....	41
4.2.4 Contexte droit	44
4.2.5 Le haut degré	49
4.3 2 ^{ème} phase – « être + ADJ + comme + SN »	52

4.3.1	Associations comparatives.....	52
4.3.2	Adjectifs « candidats » au figement.....	56
4.3.3	Locutions figées adjectivales avec <i>comme</i>	60
Conclusion	71
Résumé en français	73
Résumé český	77
Bibliographie	81
Liste des tableaux.....		85
Liste des graphiques.....		85
ANNEXE A	86
ANNEXE B.....		89
ANNEXE C.....		91

Introduction

Dans toutes les langues il arrive qu'un mot puisse, selon le contexte, faire partie de classes de mots différentes. Tel est aussi le cas du mot *comme* en français qui figure dans les dictionnaires tantôt comme un adverbe tantôt comme une conjonction. Dans mon mémoire de licence je voudrais m'interroger sur les fonctions de ce mot qui sont, quant à elles, très variées et en analyser une plus en détails. Nous allons procéder d'une analyse générale, c'est-à-dire de l'analyse à propos de tous ses emplois, jusqu'à l'analyse détaillée de l'un de ses usages, à savoir la comparaison.

Le premier chapitre est consacré à la recherche du mot *comme* dans les dictionnaires et les grammaires afin de voir quel est le classement du mot et quels emplois y figurent.

Dans le deuxième chapitre nous allons aborder *comme* dans son rôle de conjonction de subordination. Dans ce cas, *comme* peut exprimer les rapports temporels et/ou causals.

Le troisième chapitre traite de son identité de l'adverbe de manière. Deux emplois sont présentés ici, concrètement l'emploi intensif dans des phrases exclamatives et l'emploi comparatif. Ce dernier est analysé par la suite du point de vue syntaxique ainsi que sémantique. Finalement, nous allons amorcer le terme de comparaison métaphorique et les caractéristiques principales des locutions figées, qui sont des composantes obligatoires de l'étude de *comme* dans la partie empirique.

Le quatrième et le dernier chapitre qui représente la partie empirique du travail est destiné à l'étude d'une construction comparative dans laquelle *comme* est précédé du verbe *être* et d'un adjectif (*être* + ADJ + *comme*). La première phase de l'étude consiste dans la recherche de la construction dans des corpus linguistiques parallèles et monolingues, à savoir *InterCorp* du français, *L'Est républicain* et *frWaC*.

Premièrement, grâce aux résultats obtenus des corpus, nous allons examiner la fréquence de la construction dans des types de textes différents. Deuxièmement, un autre corpus, concrètement *FRANTEXT*, sera ajouté pour pouvoir comparer aussi les fréquences concernant des genres littéraires différents. Troisièmement, par le biais d'une analyse syntaxique ainsi que sémantique du contexte droit des résultats (*être* + ADJ + *comme* + X), nous tenterons d'en identifier ceux qui portent le sens figuré.

En deuxième phase de notre étude, en nous basant sur le critère fréquentiel, nous allons sélectionner des adjectifs candidats qui figureront dans des locutions figées potentielles de type « *être* + ADJ + *comme* + SN ».

Enfin, notre intérêt sera de chercher dans les dictionnaires des locutions figées avec les adjectifs choisis et d'en découvrir de nouvelles en utilisant des mesures statistiques et des tests linguistiques.

1 Comme et ses emplois

La comparaison fait sans doute partie de notre vie quotidienne. C'est pourquoi il serait intéressant de voir comment cela fonctionne du point de vue linguistique. Dans le présent travail nous allons aborder un des moyens grâce auquel nous pouvons exprimer le rapport de comparaison en français, à savoir le mot¹ *comme*. Avant de nous plonger dans une analyse détaillée de l'emploi comparatif de ce dernier, nous allons essayer tout d'abord de trouver d'autres contextes dans lesquels il peut apparaître afin de mieux comprendre son fonctionnement.

1.1 Emplois selon les dictionnaires

Pour commencer, nous allons chercher le mot *comme* dans les dictionnaires, concrètement dans *Le Nouveau Petit Robert Dictionnaire de la langue française* (version CD-ROM 2008)², *Larousse Dictionnaire de français*³ et *Le Trésor de la langue française informatisé*⁴. Ensuite, nous allons comparer ce que disent les dictionnaires choisis et comme base de la comparaison nous allons prendre quatre emplois principaux de *comme*, tels que les définissent Monneret et Rioul (1999, p. 257) dans *Questions de syntaxe française*⁵.

CONJONCTION	Cause	Temps	Exclamation	Comparaison
Le Petit Robert	✓	?		?
Larousse	✓	✓		✓
TLFi	?	?		?
ADVERBE	Cause	Temps	Exclamation	Comparaison
Le Petit Robert		?	✓	?
Larousse			✓	
TLFi	?	?	✓	?

Tableau 1.1 : Quatre emplois principaux⁶ de *comme* mentionnés dans les dictionnaires, ceux-là divisés d'après le classement parmi les adverbes ou les conjonctions (✓ = mentionné/ ? = indéterminé/ gris = classé dans l'autre groupe)

¹ Dans ce travail nous simplifions le caractère non rigoureux du terme et nous allons appeler par *mot* tout groupe de sons articulés ou graphiques.

² Appelé par la suite « Le Petit Robert »

³ Appelé par la suite « Larousse »

[En ligne] [Consulté le 17.02.2016]. Disponible sur : <<http://www.larousse.fr/dictionnaires/francais>>

⁴ Appelé par la suite « TLFi »

[En ligne] [Consulté le 17.02.2016]. Disponible sur : <<http://www.cnrtl.fr/definition/>>

⁵ Les auteurs analysent par la suite chaque emploi du point de vue syntaxique.

⁶ Emplois considérés comme principaux selon notre choix et suivant le modèle de Monneret et Rioul (1999)

Dans l'intention de mieux voir quels paramètres diffèrent ou coïncident dans les dictionnaires particuliers, nous avons élaboré un tableau. Conformément au tableau 1.1, nous pouvons constater qu'aucun dictionnaire n'a omis les fonctions probablement essentielles du mot. Ces dernières encadrent le rapport **temporel**, le rapport **causal**, la **modalité exclamative** et, en fin de compte, la **comparaison**.

Néanmoins, ce qui reste relativement ambigu en matière fonctionnelle, c'est la catégorisation du mot proposée par les dictionnaires. En effet, de la part du *Petit Robert*, on parle, à l'égard de l'emploi temporel et comparatif, de la catégorie de « Conjonctions et adverbes ». Quant au *TLFi*, il propose tout simplement le terme « morphème »⁷ sans classement explicite. En comparaison avec les deux précédents, *Larousse*, même s'il s'agit du plus bref en ce qui concerne les explications, semble avoir des concepts définis d'une manière rigoureuse.

La seule exception est l'usage de *comme* dans les phrases exclamatives, étant donné que tous les dictionnaires affirment son appartenance au groupe des adverbes.

1.2 Emplois selon les grammaires

Après avoir vu le traitement de *comme* au sein des dictionnaires, nous allons procéder maintenant de la même manière en nous penchant sur les grammaires. Nous avons élu comme référence *Francouzská mluvnice* (Hendrich, Radina & Tláskal, 2001)⁸, la *Grammaire de la Phrase Française* (Le Goffic, 1993)⁹, *Le bon usage* (Grevisse & Goosse, 2008)¹⁰ et la *Grammaire méthodique du français* (Riegel, Pellat & Rioul, 2014)¹¹.

En commençant avec l'exclamation, le tableau 1.2 confirme les résultats trouvés dans les dictionnaires puisque toutes les trois grammaires se mettent toutes d'accord sur le caractère adverbial de *comme*. Ceci est pareil pour la comparaison sauf que *Francouzská mluvnice* oscille entre les deux parties du discours. De plus, *comme* exprimant le temps et la cause est dans celle-là, contrairement à deux autres ouvrages de référence, considéré en tant que conjonction.

⁷ Le morphème est une unité minimale porteuse de sens (Dubois, 2012, p. 310). On distingue des morphèmes grammaticaux (grammèmes) et des morphèmes lexicaux (lexèmes) (Ibid.). Ainsi, nous ne pouvons pas en déduire si le mot *morphème* est envisagé ici en tant que morphème grammatical (*comme* conjonction) ou lexical (*comme* adverbe).

⁸ Grammaire descriptive du français écrite en tchèque à l'instar d'une approche contrastive des deux langues

⁹ Cet ouvrage nous offre une présentation complète des formes syntaxiques de la phrase française.

¹⁰ Grammaire descriptive du français très complexe, pleine d'explications et d'exemples surtout littéraires

¹¹ Grammaire globale du français contemporain qui traite la morphologie ainsi que la syntaxe et le fonctionnement de la langue en général

En ce qui concerne *Le bon usage*, le *comme* temporel n’y est pas au centre d’intérêt, il ne fait que partie de l’énumération de conjonctions des propositions temporelles exprimant deux faits simultanés. Cela peut paraître un peu surprenant si nous prenons en compte le caractère volumineux de cet ouvrage.

CONJONCTION	Cause	Temps	Exclamation	Comparaison
Francouzská mluvnice	✓	✓		?
Grammaire de la Phrase Française				
Le bon usage	?	✓		
Grammaire méthodique du français	✓	✓		✓
ADVERBE	Cause	Temps	Exclamation	Comparaison
Francouzská mluvnice			✓	?
Grammaire de la Phrase Française	✓	✓	✓	✓
Le bon usage	?		✓	✓
Grammaire méthodique du français			✓	

Tableau 1.2 : Quatre emplois principaux de *comme* mentionnés dans les grammaires, ceux-là divisés d’après le classement parmi les adverbes ou les conjonctions (✓ = mentionné/ ? = indéterminé/ gris = classé dans l’autre groupe)

Ce qui différencie la *Grammaire méthodique* des autres, c’est sa définition de *comme* en tant que conjonction dans son emploi comparatif. Quant à la *Grammaire de la Phrase Française*, elle se distingue par rapport à toutes les autres grammaires par son classement explicite de tous les emplois de *comme* parmi les adverbes.

Les quatre emplois indiqués dans les tableaux 1.1 et 1.2 vont être soumis à une analyse plus profonde dans le deuxième chapitre, dont un emploi va être l’objet clé de notre étude.

1.3 Difficultés du classement

Les dictionnaires et les grammaires non plus ne nous ont pas vraiment facilité la tâche. Il apparaît que la classification de *comme* n’obéisse pas à une norme bien définie. Comme le mentionnent Moline et Flaux (2008, p. 3 – 4) dans l’article *Constructions en comme : homonymie ou polysémie ? Un état de la question*, *comme* est susceptible d’apparaître dans des contextes syntaxiques très variés et, par conséquent, peut être chargé de plusieurs valeurs sémantiques. Ceci pose en réalité beaucoup de difficultés en ce qui concerne le classement du mot dans une seule partie du discours, du point de vue morphosyntaxique ainsi que sémantique. C’est la raison pour laquelle « [de] nombreux

linguistes [] voient dans *comme* tantôt un adverbe, tantôt une conjonction, voire même une préposition » (Fuchs & Le Goffic, 2008, p. 71 – 72).

Si nous prenons pour référence par exemple Fuchs et Le Goffic (2008, p. 72), « *comme*, dans tous ses emplois (subordonnants ou non), est le marqueur d'une opération portant sur le « modus »¹² (manière de faire ou d'être) et, à ce titre, il est d'essence adverbiale ». Puisque cette approche envisage le « marqueur »¹³ qui renvoie « à ce qui est sous-jacent à tous les types d'emplois », elle peut être considérée comme polysémique (Portine, 1995, p. 369). En revanche, il existe des hypothèses selon lesquelles *comme* employé différemment selon le contexte est le résultat d'une homophonie. Cela signifie que « [le] marqueur peut renvoyer à un type d'emplois » (Ibid.)¹⁴.

Le fait de juger *comme* de nature purement adverbiale peut être lié à l'étymologie de celui-ci. En effet, le mot tel que nous le connaissons aujourd'hui vient à l'origine de « *quomodo* » latin qui veut dire « de quelle manière » (Grevisse & Goosse, 2008, p. 1385).

En résumé, il semble qu'il soit difficile de classer *comme* d'une manière univoque de telle sorte que cela provoque assez de controverses dans le domaine de la linguistique.

1.4 Emplois secondaires

Dans le but de rendre complet le schéma des emplois de *comme*, les emplois secondaires figurant dans les dictionnaires et les grammaires vont être mentionnés brièvement¹⁵.

- **ADDITION :** Nous pouvons nous servir de *comme* non seulement dans des phrases subordonnées, mais aussi pour exprimer l'addition dans des phrases coordonnées.

(1) « *Sur la terre **comme** au ciel* » (Le Petit Robert)

Dans ce cas, *comme* peut substituer des conjonctions de coordination *ainsi que* ou tout simplement *et*.

- **ATTRIBUTION :** Il est assez commun d'utiliser *comme* dans le but d'attribuer une qualité ou définir « l'appartenance à une classe » (Portine, 1995, p. 373).

¹² Fuchs préfère de parler de *modus* plutôt que de *manière* à cause de la signification inexacte de celle-ci.

¹³ Dans ce cas il est question du marqueur sémantique.

¹⁴ Les auteurs qui se distinguent par rapport à la conception polysémique et homophonique de *comme* sont cités et la problématique est traitée en profondeur par Moline et Flaux (2008).

¹⁵ Dans l'**Annexe A** de ce travail nous pouvons voir la répartition de chacun de ces emplois dans les dictionnaires ainsi que les grammaires traitées précédemment selon le classement parmi les adverbes ou les conjonctions. Si la mention d'un emploi n'a pas été trouvée, cela veut dire que la source référentielle ne lui consacre pas l'espace dans un chapitre ou un paragraphe individuel. Néanmoins, il n'est pas exclu que *comme* y soit présent sous forme d'un exemple ou seulement faisant partie d'une énumération dans un type d'emploi.

(2) « *Elle a été choisie **comme secrétaire.*** » (Hendrich, Radina & Tláskal, 2001, p. 538).

Au niveau syntaxique, *comme* introduit souvent des attributs (Ibid.).

- **APPROXIMATION** : Le fait que *comme* contribue à l'expression de l'approximation peut avoir des rapports avec son caractère indéfini.

(3) « *Il entendit **comme un bruit*** » (Moline & Flaux, 2008, p. 4)

En effet, Le Goffic (1993, p. 391) considère *comme* en tant que « intégratif » qui répond à l'interrogatif *comment*, et il est d'avis que l'indéfinition propre à ce dernier reste conservée aussi dans la réponse avec *comme*. Pour lui, tous les adverbes connecteurs portent seulement sur une relation abstraite et donc indéfinie, que ce soit une spécification de temps, de lieu ou de manière (op. cit., p. 392).

- **EXEMPLIFICATION** : Le *TLFi* nous procure l'explication suivante : « [*Comme*] sert à introduire un ex. [sic] qu'il présente comme conforme (ou quasi-identique) au modèle implicitement envisagé ».

(4) « *Plus les sciences se rapportent à l'homme, **comme la médecine par exemple**, moins elles peuvent se passer de religion [...] »¹⁶*

- **INTERROGATION INDIRECTE** : *Comme* est parfois considéré comme équivalent de l'adverbe interrogatif *comment* dans l'interrogation indirecte et il est ainsi utilisé surtout dans la langue littéraire (Grevisse & Goosse, 2008, p. 1218).

(5) « *Sais-tu **COMME** [c'est l'auteur qui utilise les capitales] l'on pend ? »¹⁷*

Il était même commun de l'utiliser dans les interrogations directes jusque vers le milieu du XVII^e siècle (Ibid.).

¹⁶ J. DE MAISTRE, *Les Soirées de Saint-Petersbourg*, t. 2, 1821, p. 217 ; cité dans *TLFi*

¹⁷ CAMUS, Justes, Pl., p. 386 ; cité dans Grevisse et Goosse (2008, p. 1218)

2 *Comme* conjonction de subordination

Malgré les difficultés que le classement du mot étudié peut nous poser, pour le reste de ce travail, nous allons envisager que d'un côté, dans son emploi temporel et causal il incline plutôt vers le groupe des conjonctions et de l'autre côté, étant le moyen pour exprimer la comparaison et l'exclamation il est de nature davantage adverbiale. Le présent chapitre va être consacré à son « identité » de conjonction.

Comme fait sans doute partie des conjonctions les plus fréquentes parmi les subordonnants dans les **subordonnées circonstancielles**¹. Si la classe des adverbes est nommée parfois la classe « poubelle », parmi les circonstancielles sont classées toutes les propositions qui ne sont ni des relatives ni des complétives (Riegel, Pellat & Rioul, 2014, p. 841).

En comparaison avec des locutions conjonctives composées, morphologiquement il s'agit d'une forme simple de conjonction. Ceci n'exclut pas bien sûr la possibilité de combinaison avec une autre conjonction. D'où l'existence de nombreuses locutions conjonctives avec *comme* ayant leurs significations à part.

2.1 *Comme* conjonction de temps

Avant tout, pour ce qui est de l'enchaînement thématique, en tant que porteur du thème, *comme* doit être placé obligatoirement en tête de phrase. Le Goffic (1993, p. 483) lui attribue ici la valeur de « **l'adverbe de phrase** ».

La raison pour laquelle nous optons quelquefois pour le *comme* temporel est d'exprimer la simultanéité de deux actions au passé (Hendrich, Radina & Tláskal, 2001, p. 538).

(6) « **Comme nous sortions du bois, il a commencé à neiger.** » (op. cit., p. 533)²

Si d'autres conjonctions temporelles sont transposables dans tous les temps, *comme* est difficilement compatible avec le passé simple ou le futur (Riegel, Pellat & Rioul, 2014, p. 849). Par contre, comme dans l'exemple (6), c'est l'imparfait qui l'accompagne dans la majorité des cas à cause de l'aspect sécant de celui-là (Monneret & Rioul, 1999, p. 268). Et puisque l'imparfait est le temps issu de l'arrière-plan, *comme* participe lui aussi à l'addition

¹ Grevisse et Goosse (2008, p. 1475) utilisent le terme des propositions « adverbiales ».

² *Comme* peut être remplacé ici par la conjonction *tandis que* (Hendrich, Radina & Tláskal, 2001, p. 533).

des renseignements complémentaires sur le décor et les circonstances de l'action principale (Monneret & Rioul, 1999, p. 268).

Or, il faut souligner que, contrairement à *quand*, la présence de l'imparfait n'inclut pas le sens itératif (Hendrich, Radina & Tláskal, 2001, p. 538). D'ailleurs, Le Goffic (1993, p. 483) est d'avis qu'il existe encore une autre nuance distinctive entre *comme* et *quand* qui est la suivante: « [...] la concomitance marquée par *comme* est moins strictement et chronologiquement temporelle qu'elle n'est avec *quand* [...], elle marque plutôt une circonstance d'accompagnement de P2 [proposition 2] qu'un renseignement d'horloge (une simple indication de simultanéité). ».

2.2 *Comme* connecteur de cause

La Grammaire méthodique (Riegel, Pellat & Rioul, 2014, p. 1044) désigne des connecteurs comme termes de liaison assurant la cohérence du texte « en marquant des relations entre les propositions ou entre les séquences qui [le] composent ». Quant au connecteur étudié, il appartient à l'ensemble des connecteurs argumentatifs qui enveloppent le rapport de causalité (Grevisse & Goosse, 2008, p. 1489).

Pour relier à ce dont nous avons parlé dans le sous-chapitre précédent, la valeur causale de *comme* ne doit pas être explicitement séparée de la valeur temporelle. En réalité, quelques linguistes la conçoivent seulement comme un sous-ensemble de celle-ci, car ce ne sont que des éléments extralinguistiques qui impliquent le sens causal d'un énoncé (Moline, 2008, p. 88). Ce que partagent le *comme* causal et le *comme* temporel, c'est entre autres, l'introduction de la proposition thématique. Evidemment, dans la relation logique cause-conséquence, la cause précède cette dernière (Loucká, 2010, p. 28).

Par ailleurs, en tant que connecteur de cause, *comme* peut être remplacé par son équivalent *puisque*. Or, il existe des nuances qui déterminent l'usage de l'un ou de l'autre dans un contexte spécifique. Si *comme* est placé strictement au début d'une proposition, *puisque* ne doit pas respecter cette contrainte syntaxique de telle manière qu'il est fréquemment sujet à la postposition (Monneret & Rioul, 1999, p. 269). Pourtant, en fonction attributive, *comme* peut être postposé à une épithète ou une apposition détachée tout en gardant le rapport de cause (Grevisse & Goosse, 2008, p. 1493)³.

³ **Exemple :** « *Faible comme elle est, j'ai peur de la voir tomber malade à chaque Carême* »
BOURGET, *Drames de famille*, p. 19 ; cité dans Grevisse et Goosse (2008, p. 1493)

Il existe aussi une nuance sémantique, et c'est que *puisque* comporte en lui une certaine présupposition du locuteur que la cause annoncée est au moment de l'énonciation déjà bien connue (Portine, 1995, p. 396).

(7) « **Comme/Puisqu'il a neigé toute la nuit, la route est dangereuse.** » (Loucká, 2010, p. 28)

La phrase illustrative indiquée en (7) peut être saisie de deux perspectives. L'une est propre à l'interlocuteur pour lequel le contenu propositionnel était jusqu'ici inconnu et l'autre vient d'un destinataire conscient du fait.

En guise de conclusion, ceci n'était qu'une brève présentation de *comme* dans ses emplois conjonctifs. Plus d'espace sera dédié à ses emplois adverbiaux dans le chapitre 3 avec l'accent mis sur la comparaison.

3 *Comme* adverbe de manière

Parmi les propositions circonstancielles on distingue sept espèces, nommément les propositions de temps, de cause, de manière, de conséquence, de but, de concession et de condition (Grevisse & Goosse, 2008, p. 1479). Bien entendu, ce sont les propositions adverbiales de manière que *comme* introduit.

Comme nous l'avons déjà esquissé dans 2.1, le sens du mot est lié également à la position de la subordonnée dans laquelle il apparaît par rapport au verbe principal (Le Goffic, 1993, p. 482). Ainsi, son incidence syntaxique est double.

En tant qu'adverbe de phrase, *comme* est porteur de plusieurs significations examinées antérieurement et dans cette position, il est censé figurer en tant que conjonction. Il nous reste un sens non encore mentionné que nous avons décidé d'insérer dans le chapitre concernant les adverbes. En effet, en tant qu'élément introducteur, *comme* est l'un des moyens utilisé pour définir le type d'une phrase. Dans ce cas, nous faisons allusion à des phrases exclamatives, sujet analysé dans le premier chapitre de cette partie du travail.

En revanche, en position postverbale, Le Goffic (Ibid.) propose la dénomination de « **l'adverbe de prédicat** ». Cela signifie que la portée de l'adverbe ne s'effectue que sur le niveau intra-prédicatif. Dans cette optique nous allons aborder enfin et pour le reste du travail étudié, l'emploi prototypique de *comme* qu'est la comparaison.

3.1 Exclamation d'intensité

La Grammaire méthodique (Riegel, Pellat & Rioul, 2014, p. 683) nous fournit cette caractéristique des énoncés exclamatifs : « Ils expriment l'affectivité, un sentiment plus ou moins vif du locuteur à l'égard du contenu de son énoncé et ils jouent un rôle important dans la communication orale ». Ce qui indique un tel type de phrase c'est principalement le point d'exclamation¹ à l'écrit et l'intonation très contrastée à l'oral (Grevisse & Goosse, 2008, p. 505). Sinon, l'ordre des mots reste comme dans une phrase déclarative (Hendrich, Radina & Tláskal, 2001, p. 611).

Sur le plan sémantique, le marqueur de la modalité exclamative², qu'est dans notre cas *comme*, signale un « haut degré dans l'ordre de la quantité ou de la qualité »³. En

¹ Le point d'exclamation est délibérément omis par certains auteurs même s'il est clair que la phrase est de nature exclamative (Grevisse & Goosse, 2008).

² La modalité exclamative est appelée des fois « interjective ».

³ J.-C. Milner, 1978, p. 252 ; cité dans Riegel, Rioul et Pellat, 2014, p. 687

général, les éléments auxquels se rattache l’adverbe exclamatif sont séparés de celui-ci (Grevisse & Goosse, 2008, p. 506), ce qui est observable dans l’exemple ci-dessous.

(8) « *Comme tu es sûr de toi! (...) Comme tu es fort! (Anouilh)* » (Riegel, Rioul & Pellat, 2014, p. 688)

Notons que dans la proposition autonome, l’adverbe peut se substituer par *combien* ou par *quel/ce quel/qu’est-ce que/comment que* qui sont, quant à eux, propres à la langue familière voire populaire (Grevisse & Goosse, 2008, p. 1251).

En outre, « [d]e la même façon qu’on distingue une interrogation indirecte, on doit parler d’une exclamation indirecte » (Grevisse & Goosse, 2008, p. 504). Par conséquent, nous pouvons tomber sur *comme* pourvu de sens exclamatif aussi au niveau de ce type de phrases.

(9) « *Regarde comme il est sage.* » (op. cit., p. 1528)

Comme nous pouvons le voir en haut, les verbes dits « supports » liés à cet usage sont notamment les verbes d’opinion ou de perception, employés souvent à l’impératif (Ibid.).

3.2 Comparaison

Pour commencer, pour ce qui touche l’étymologie du mot *comparaison* lui-même, en latin ce terme, composé de la préposition *cum* (= avec, ensemble) et de l’adjectif *par*, *paris* (= égal, pareil) désignait « l’action de comparer » ce qui s’est, bien évidemment, sauvegardé jusqu’au français contemporain (Fuchs, 2014, p. 11).

Nous pouvons nous poser maintenant la question de savoir ce que cela veut dire exactement. Conformément à la définition du *TLFi*⁴, cela signifie faire un acte mental pour pouvoir confronter deux ou plusieurs éléments de nature quelconque, mais à condition que ceux-là soient réciproquement comparables, à l’instar de leurs marques de ressemblance ou différence.

Les subordonnées comparatives sont d’habitude classées sous les subordonnées circonstancielles, bien qu’en se rendant compte de leurs spécificités, quelques linguistes commencent actuellement à les mettre à part (Fuchs, 2014, p. 16).

⁴ [En ligne] [Consulté le 09.03.2016]. Disponible sur : <<http://www.cnrtl.fr/definition/>>

3.2.1 Types de comparaison

Pour comparer deux entités, nous pouvons adopter la terminologie introduite par Fuchs (2014)⁵. Concrètement, il est question de deux « **comparandes** » dont l'une est le « **comparé** »⁶, et l'autre le « **comparant** »⁷, mises en relation par l'intermédiaire d'une propriété spécifique, appelée « **paramètre** » (op. cit., p. 22). Il nous reste encore de nommer, d'un côté les « **marqueurs du paramètre** » qui signalent le degré et, de l'autre côté les « **marqueurs du comparant** » qui servent à introduire celui-ci (Ibid.).

- (10) « *Pierre* [comparé] *est plus (ou aussi)* [marqueur du paramètre] *grand*
[paramètre *être grand*] *que* [marqueur du comparant] *Paul* [comparant]. »
(op. cit., p. 23)

Comme le présente Fuchs (op. cit., p. 24) dans la phrase illustrative ci-dessus, le schéma traditionnel de la comparaison est présenté par la corrélation de deux constituants, à savoir d'un adverbe de degré et de marqueur *que*.

Or, le français a à sa disposition une échelle beaucoup plus étendue en ce qui concerne l'expression du sens comparatif. En règle générale, on distingue le **degré absolu** ou « superlatif absolu » et **relatif** (Grevisse & Goosse, 2008, p. 1221). Ce dernier peut former le 1^{er} degré, appelé « le comparatif », et le 2nd degré, appelé le « superlatif relatif » (Ibid.).

La comparaison de deux éléments se réalise au moyen d'une quantification numérique ou d'une quantification évaluative (Fuchs, 2014, p. 40). La quantification numérique correspond à une gradation objective implicite, fondée sur les paramètres objectivement « mesurables », et l'autre est le fruit d'une indication subjective explicite lorsque la mesure de la quantité n'est pas spécifiée (Ibid.).

La Grammaire de la Phrase Française (Le Goffic, 1993, p. 402) utilise la terminologie de comparaison « **quantitative** » avec l'antécédent adverbial et à l'inverse, de la comparative « **qualitative** » dont l'antécédent est un adjectif⁸.

⁵ Fuchs (2014, p. 17) explique que jusque-là, l'expression de la comparaison en français n'a pas été développée de manière cohérente et intégrale.

⁶ Parfois nommé « cible » dans l'opposition cible/standard (Desmets, 2008, p. 34)

⁷ Fuchs (2014, p. 22) décide d'appeler cette entité « standard ».

⁸ Desmets (2008, p. 33) utilise les termes de comparaison scalaire (quantitative) et non-scalaire (qualitative).

3.2.2 Comparaison avec *comme*

Comme amorcé déjà au début de ce travail, *comme* participe lui aussi à l'expression de la comparaison, ce qui est assurément son emploi le plus fréquent (Monneret & Rioul, 1999, p. 258). Dans ce cas-là, il s'agit de ce qu'on appelle la comparaison « **similative** » ou la comparaison « **de ressemblance** » (Fuchs 2014, p. 133). Voilà l'explication de Fuchs (Ibid.) : « C'est une comparaison qualitative qui n'opère pas de gradation quantitative sur un paramètre, mais rapproche les deux comparandes sur la base d'une similarité ou d'une analogie ». La similarité consiste à comparer des entités homogènes tandis que l'analogie ressort des entités hétérogènes (op. cit., p. 134).

Si nous remplacions la configuration corrélatrice *aussi...que* dans l'exemple canonique « *Pierre est aussi grand que Paul* » par *comme*, nous pourrions repérer une nuance distinctive. En effet, alors que la structure comparative avec *aussi...que* requiert au moins un minimum de grandeur pour le comparé, avec *comme* la notion de grandeur ne nous dit rien sur la taille ni du comparé ni du comparant (Portine, 1995, p. 374). Tout ce que nous savons est seulement que le comparé n'est ni plus ni moins grand que le comparant (op. cit., p. 375).

Observons maintenant les exemples présentés par Fuchs (2014, p. 138) :

(11) « *Paul travaille **autant qu'**un vrai professionnel.* »

(12) « *Paul travaille **comme** un vrai professionnel.* »

Si nous appliquons *que* ou *comme* dans le même type de phrase, nous nous apercevons de ce trait distinctif qu'est la quantité (11)/qualité (12). Dans (11) la quantité de travail de Paul est égale à la quantité de travail d'un vrai professionnel. Par contre, dans (12) la manière de travail de Paul est identifiée à la manière de travail d'un vrai professionnel (Ibid.).

Néanmoins, Moline (2006, p. 15) explique que *comme* est compatible même avec une forme de manière quantifiante, ce qui est contraint par le type du verbe utilisé. Elle oppose les verbes tels que *se comporter* et *pleuvoir*, dont l'un n'admet que la signification non quantifiante et l'autre que quantifiante (Ibid.).

D'ailleurs, ce qui différencie encore les morphèmes *que* et *comme*, c'est que *comme* n'a besoin d'aucun antécédent, c'est-à-dire qu'il ne requiert pas la corrélation obligatoire avec un marqueur du paramètre (*autant* dans l'exemple (11)) (Fuchs, 2014, p. 138).

En outre, *comme* joue non seulement le rôle de « **modifieur** », mais lui aussi a ses propres modifieurs qui le précèdent. Les uns permettent de mettre en relief une stricte

identité de manière, par exemple *tout, tout à fait, exactement*, etc., les autres servent à marquer une nuance d'approche, par exemple *presque, à peu près*, etc. (Fuchs, 2014, p. 138)⁹.

3.2.3 *Comme* + X

D'après Fuchs (2014, p. 146 – 150) « la comparaison de ressemblance en *comme* est une **identité de manière** [...] ». Dans une proposition, *comme* peut être suivi de divers contextes syntaxiques. Pour comprendre mieux le phénomène, nous allons parcourir désormais toutes les possibilités :

➤ SYNTAGME VERBAL (SV)

(13) « *On n'écrit pas l'histoire en 2012 **comme** on l'écrivait en 1950 (Le Monde des livres)* » (op. cit., p. 139)

Si nous observons l'exemple (13), nous pouvons constater que le verbe dans la subordonnée est anaphoriquement repris de la proposition « matrice » (Ibid.). Sans compter que le verbe « plein » peut être remplacé même par un verbe différent, en général par un verbe support comme *être* ou *faire* (Ibid.) :

(14) « *Peut-on **aimer** Dieu **comme** vous **faites** et **détester** si cruellement ses œuvres? (Sand)* » (op. cit., p. 140)

Etant donné la propriété de *comme* à relier deux propositions, Monneret et Rioul (1999, p. 259) lui attribuent ici le rôle de conjonction.

Du côté sémantique, il s'agit de ce que Fuchs (2014, p. 146) appelle « **l'identité de manière de faire** » entre le comparé et le comparant.

Par ailleurs, il existe aussi un type de comparaison métalinguistique basée sur la **ressemblance entre énonciations** (op. cit., p. 158).

(15) « *Et puis, **comme dit** cet autre, le fusil ne peut pas être toujours tendu (Lavedan)* » (Ibid.).

Le verbe dans la subordonnée est le plus souvent le verbe *dire* ou ses variantes (Ibid.).

➤ SYNTAGME NOMINAL (SN)/son substitut pronominal

⁹ **Exemple** : « *Les choses se sont passées **exactement comme** il avait prévu* » (Fuchs, 2014, p. 138). D'ailleurs, il faut dire que le *comme* temporel ou causal présentés dans le chapitre 2 sont incompatibles avec ces marqueurs cités (Moline, 2008, p. 95).

Si le prédicat entre les deux comparandes est un prédicat attributif ou seulement le verbe copule, la relation comparative peut être envisagée comme « **identité de manière d'être** » (Fuchs, 2014, p. 148). Avec le prédicat attributif, la propriété est désignée par un adjectif et le comparant peut être soit **générique** exprimant le haut degré soit **spécifique** selon le contexte.

D'après Fuchs, nous pouvons rencontrer aussi des phrases qui sont à mi-chemin entre le prédicat attributif et la copule, lorsque la propriété en jeu n'est pas exprimée, mais est facilement déductible, comme par exemple dans « [...] *ses bras étaient comme de la neige* », l'adjectif *blancs* reste implicite (op. cit., p. 149).

Puis, une des spécificités de *comme* est de former des structures elliptiques. *Le bon usage* (Grevisse & Goosse, 2008, p. 226) présente l'**ellipse** en tant que l'absence d'un ou de plusieurs mots avec lesquels on exprimerait une idée dans son intégralité. Cela veut dire que dans la subordonnée introduite par *comme*, le prédicat de la principale n'est pas répété.

(16) « *Il est malin comme un singe.* » (Monneret & Rioul, 1999, p. 259)

L'exemple indiqué nous procure une construction elliptique par excellence. En effet, la phrase est complète, pour peu qu'elle soit restituée en « *Il est malin comme un singe est malin* » (Ibid.). Le comparé *il* et le comparant *un singe* sont sur le même plan syntaxique (op. cit., p. 260). En d'autres mots, les deux accomplissent la fonction de sujet et partagent le prédicat commun *être malin* (Ibid.). *Comme* est dans ce type de construction assimilé parfois à une préposition (op. cit., p. 259). Nous voyons donc qu'avec le verbe *être* et le SN sans expansion, la reconstitution de la phrase est en général bien transparente.

Cependant, le caractère ambigu de la construction peut être visible dans des phrases comportant des verbes autres que copules et des compléments. Observons une phrase exemplaire :

(17) « *Elle traite sa fille comme son mari.* » (Fuchs, 2014, p. 140)

Celle-ci veut dire soit « elle traite sa fille comme elle traite son mari » soit « elle traite sa fille comme son mari traite sa fille »¹⁰.

➤ CIRCONSTANT (par ex. adverbe)

(18) « *Il la vénère exactement comme autrefois.* » (Ibid.)

¹⁰ Une autre ambiguïté pourrait survenir si la phrase contenait le prédicat composé de plus d'un verbe à sens plein encore un verbe modal. Ainsi, il ne serait pas clair si la modalité devrait être reprise aussi.

Dans cette phrase illustrative nous faisons face à une construction dans laquelle une entité (il) est dédoublée en deux (*il d'aujourd'hui* et *il d'autrefois*) (Fuchs, 2014, p. 150). Fuchs (Ibid.) remarque que ce type de comparaison est souvent interprété comme une « **ressemblance entre situations** ».

➤ ADJECTIF, PARTICIPE PASSE/PRESENT

(19) « *Elle semblait comme pétrifiée.* » (op. cit., p. 140)

La phrase plus haut nous fait penser au sens **approximatif** de *comme* (cf. 1.4). En effet, le comparé (elle) sert de support à une prédication attributive non spécifiée, identifiée à une autre prédication attributive, cette dernière représentant en quelque sorte le comparant (op. cit., p. 161 – 162). Nous avons donc affaire à une ressemblance vague entre deux prédications attributives (op. cit., p. 162).

➤ COMBINAISON DE CONSTITUANTS (par ex. SN + circonstant)

(20) « *Il s'ennuie en ville comme moi à la campagne.* » (op. cit., p. 140)

Cette phrase est formée de deux entités (*il* et *moi*) et d'une « variable » spatiale (op. cit., p. 150). De nouveau, il est question d'une ressemblance entre situations (Ibid.)¹¹.

3.2.4 Comparaison métaphorique

Jusqu'ici nous avons présenté principalement des exemples issus de la comparaison dite « **courante** », dans lesquels deux comparandes appartaient au même champ de pensée (Kokelberg, 1993, p. 69). Désormais il est temps de découvrir également des comparaisons nommées « **littéraires** »¹² qui, en associant des réalités qui n'ont à priori rien en commun, ouvrent le passage à l'imagination (op. cit., p. 70).

Tout d'abord, une des propriétés de *comme* est d'exprimer le **haut degré**¹³ (superlatif absolu) en mettant côte à côte une entité spécifique et un « **comparant-parangon** » générique qui ne doit pas se limiter à un seul (Fuchs, 2014, p. 160). Ainsi, nous avons plusieurs possibilités pour exprimer le fait de beaucoup travailler, par exemple « *travailler comme un bœuf* » ou « *travailler comme un fou* » (Ibid.). Sans connaissance

¹¹ En effet, *comme* dans cet exemple qualifie la manière de s'ennuyer. En revanche, si l'on remplaçait la subordonnée par une subordonnée détachée avec une virgule, *comme* aurait une portée extra-prédicative et le paramètre à comparer serait plutôt le fait même de s'ennuyer (Fuchs, 2014, p. 153).

¹² Une autre terminologie : comparaison à sens propre (logique)/figuré (tropologique)

¹³ Ce que partagent le *comme* exclamatif et le *comme* comparatif, c'est justement la sémantique intensive de haut degré (cf. 3.1).

même du comparant, il est possible de déduire la signification de haut degré (Grevisse & Goosse, 2008, p. 1243).

Dans la tradition rhétorique l'on parle du **sens figuré** lorsqu'on a un **trope** (ou figure de sens), c'est-à-dire un détournement du sens conventionnel ou littéral (Riegel, Pellat & Rioul, 2014, p. 934). La plupart des sens figurés sont issus de la **métaphore** (Dubois, 2012, p. 302). La métaphore est en effet une sorte de « **comparaison figurative** » basée sur le principe de l'analogie (Fuchs, 2014, p. 165) ou, à l'inverse, une comparaison fondée sur une démarche métaphorique peut être nommée « **comparaison métaphorique** » (Kokelberg, 1993, p. 87). Il se pose donc la question de savoir quelle est la différence entre une comparaison et une métaphore.

Premièrement, Portine (1995, p. 377) évoque dans son article que tandis que la comparaison permet d'évaluer un jugement, la métaphore opère sur la double référenciation (Ibid.). Effectivement, si la comparaison laisse les réalités comparées à une certaine distance, dans la métaphore les deux réalités sont quasiment identifiées (Kokelberg, 1993). Le fonctionnement de la métaphore est tel qu'elle assimile le comparé au comparant sans outil de comparaison (Fuchs, 2014, p. 165).

(21) « *Paul est ignorant comme un âne* » vs. « *Paul est un âne.* » (Dubois, 2012, p. 302)

Deuxièmement, dans la mesure où la comparaison de haut degré est déjà socialement codée voire codifiée dans les dictionnaires, nous pouvons tomber sur des expressions poncives, c'est-à-dire les « **clichés** » (Kokelberg, 1993, p. 71). De même, quant à la métaphore, si son usage est déjà figé, il faut parler plutôt de **catachrèse**. Bien entendu, une métaphore doit rester originale, sinon en la répétant voire en la lexicalisant elle perd de sa force (Riegel, Pellat & Rioul, 2014, p. 936).

Troisièmement, l'art d'une expression métaphorique est selon Kokelberg (1993, p. 91 – 92) de provoquer une réaction immédiate sur l'imagination en suggérant une association. Le lecteur est de cette manière invité à dépasser le sens propre d'un mot pour chercher le sens figuré (op. cit., p. 92). D'où les lectures plurielles possibles, entraînées par des **connotations** différentes (op. cit., p. 94). Par contre, la comparaison qui, par une **dénotation** directe facilite la compréhension du lecteur, favorise la lecture sur la base « rationnelle » (op. cit., p. 86 et 91).

Fuchs (2014, p. 13, 15) développe dans son ouvrage une idée intéressante à propos de la comparaison omniprésente dans le domaine de la publicité, découlant de la

concurrence. En effet, le langage figuratif dans le domaine du journalisme est basé sur le principe de transfert des sens connotatifs d'un contexte vers un autre contexte, dans lesquels ils partagent au moins un **sème**, c'est-à-dire un trait sémantique commun (Čechová, Krčmová & Minářová, 2008, p. 251). Bien entendu, en comparaison avec des locutions littéraires, les locutions figuratives y doivent être toujours sémantiquement transparentes sans quoi l'information à transmettre ne serait pas déchiffrable (op. cit., p. 252).

Dans la partie suivante nous allons identifier les caractéristiques primordiales des locutions figées qui portent, quant à elles, des connotations bien enracinées.

3.2.5 Locutions figées

Si l'on consulte le terme « **figement** » dans *le Dictionnaire de linguistique* (Dubois, 2012, p. 202), on y trouve la définition suivante : « Le figement est le processus linguistique par lequel un groupe de mots dont les éléments sont libres devient une expression dont les éléments sont indissociables. ». Il en résulte que pour pouvoir parler du figement, il faut qu'une séquence soit polylexicale, c'est-à-dire composée de deux ou plus unités lexicales autonomes (Gross, 1996, p. 9).

Gross (1996) souligne que ce qui caractérise une « **locution figée** »¹⁴, c'est primo son opacité sémantique, secundo la structure non-transformable et tertio la non-actualisation de ses composants individuels dans le discours.

D'abord, s'il est possible de déduire le sens d'une phrase à partir des éléments qui la composent, on dit que la phrase a le sens compositionnel (Gross, 1996, p. 10). Et la « compositionnalité » est justement la propriété qui manque à des locutions figées (op. cit., p. 10 – 11). Apparemment, sans connaître la sémantique d'une locution comme tout, séparément les mots composants ne nous renseignent guère. En effet, les sources du figement peuvent se rattacher à l'histoire, la mythologie, la religion ou à la littérature (Gross, 1996, p. 21 – 22). Pourtant, l'origine de certaines expressions comme par exemple *fumer comme un pompier* reste sémantiquement opaque (Moline, 2006, p. 18).

Ensuite, *Le Dictionnaire de linguistique* (Dubois, 2012, p. 202) définit une locution figée comme la suite de mots qui n'admet pas du tout ou, le cas échéant, seulement des variations des articles, des temps verbaux et des adjectifs insérés. Evidemment, tout

¹⁴ Nous envisageons une locution figée comme l'équivalent de « collocation » en anglais. Une locution qui est figée et en plus pourvue du sens idiomatique est appelée « **l'expression idiomatique** » ou « **idiotisme** ». Le mot « idiome », ayant en français pour signification « langue propre à une communauté » (TLFi), se confond parfois avec « idiom » en anglais.

dépend de la locution concrète et de son « **degré de figement** ». Etant donné que toutes les locutions ne sont pas figées au même degré, il n'est pas toujours évident jusqu'à quel endroit d'une séquence le figement ait son impact (Gross, 1996, p. 15).

Le cas le plus simple est quand on a des séquences dont l'ensemble est entièrement figé (Ibid.). Les séquences pareilles peuvent être des phrases entières appelées « proverbes » ou encore des suites verbales, nominales, adjectivales et prépositives (Ibid.).

Enfin, l'absence d'actualisation est démontrée dans la phrase suivante :

(22) « *Marie est à (la, *une, *cette, *notre) mode.* » (op. cit., p. 14).

Si la phrase (22) veut conserver son sens tel que nous le connaissons, l'article défini dans l'adjectif composé *à la mode* ne peut être substitué par aucun autre déterminant. Il s'agit donc d'un « figement total » (op. cit., p. 19).

Bref, en règle générale, une locution figée n'admet aucune modification syntaxique, aucune insertion des éléments nouveaux et ne forme pas de paradigmes (Gross, 1996).

Pourtant, par des procédés métaphoriques en littérature ou humoristiques en presse, une locution figée peut subir des transformations paradigmatiques (op. cit., p. 20). C'est ce qu'on appelle le « **défigement** ». Le but en est simple, en publicité ce jeu de langage représente une des techniques pour attirer l'attention, en littérature cela apporte de l'originalité (op. cit., p. 19 – 20). L'effet de surprise chez le lecteur entraîne, paradoxalement, l'existence sûre d'une locution figée bien connue (op. cit., p. 20).

Le chapitre 4.3 dans la deuxième partie de ce travail sera consacré à l'étude des locutions figées adjectivales avec *comme*, dont la fréquence et le degré de figement sera l'un de nos objets d'intérêt.

4 Construction « être + ADJ + comme »

Dans les pages que nous venons de voir, tous les emplois du mot *comme* ont été présentés. Nous rappelons que la comparaison est probablement son emploi le plus fréquent, et il est à supposer qu'il crée beaucoup de structures comparatives différentes (cf. 3.2.2). Même s'il serait intéressant de les voir toutes, le volume de ce travail ne nous le permet pas. Aussi avons-nous décidé de faire l'analyse d'une construction comparative plus restreinte dans laquelle *comme* soit précédé du verbe *être* et d'un adjectif¹. Notre étude est donc consacrée uniquement à « l'identité de manière d'être » (cf. 3.2.3). Fuchs (2014, p. 160) affirme que la structure la plus courante pour exprimer le **haut degré** est justement celle avec le prédicat attributif *être + adjectif* et avec **un comparant générique** (cf. 3.2.4).

Ainsi, les structures qui ne vont pas être prises en compte sont par exemple :

- séquences exprimant « l'identité de manière de faire » (cf. 3.2.3) avec un verbe quelconque, autre que copule : « *Il parle comme une vache espagnole* » (op. cit., p. 147)
- phrases avec la copule implicite : « *une fois j'arrivai le soleil comme une orange [...]* (Simon) » (op. cit., p. 149)
- phrases avec des incises : « *Quoique je sois armé, sûr de mon coup, comme un homme qui a jadis chassé le tigre [...]* (Balzac) » (op. cit., p. 154)
- schéma « ADJ + *comme* + prop. attributive » : « *Intelligent comme il l'est, je le crois en route pour la gloire.* »².
- épithètes détachées : « *Les christs comme éclairés d'une blanche lumière* »³

4.1 Hypothèse et méthodologie

Tout en considérant le caractère comparatif de **similitude** de la construction (cf. 3.2.2), nous estimons que la construction est présente dans n'importe quel type de discours. Néanmoins, nous présupposons que sa fréquence ne soit pas équilibrée dans des

¹ Nous supposons qu'il s'agisse principalement des adjectifs qualificatifs, d'un côté parce que ces derniers nous intéressent dans ce travail, de l'autre côté parce que les adjectifs relationnels se prêtent moins fréquemment à former des structures attributives.

² GIDE, *Les Faux-monnayeurs*, 1925, p. 1117 ; cité dans *TLFi*
Néanmoins, la séquence qui correspond à notre schéma est par exemple « [...] *elle n'aurait pas été malheureuse comme elle l'est* [...] » (FRANTEXT, Bosco, H., *Le Mas Théotisme*, 1945).

³ A. POMMIER, *Océanides et fantaisies*, 1839, p. 122; cité dans *TLFi*

types de textes différents et que ces derniers déterminent l'usage des **comparaisons courantes** ou **littéraires** (cf. 3.2.4).

D'abord, nous émettons l'hypothèse que dans les **textes littéraires** la construction examinée soit très fréquente, car la **comparaison métaphorique** fait partie des figures de style assez recherchées par les écrivains, notamment par les poètes. Selon Fuchs (2014, p. 157) « *comme* [préserve] des possibilités d'interprétation multiples, avec un parfum ou un halo d'indétermination propice au déploiement de la sensibilité du lecteur, en écho à celle du poète. ». Par conséquent, la suite pourrait refléter une plus grande créativité de l'auteur, ce qui pourrait donner naissance à des comparaisons inédites.

Ensuite, en matière fréquentielle, cela pourrait être valable même pour les **textes tirés du web** et les **textes journalistiques**. En comparaison avec les textes littéraires, les locutions y pourraient être beaucoup plus figées voire **clichées** (cf. 3.2.4). Il est assez probable que dans les deux types de textes nous pourrions rencontrer des exemples de **défigement** (cf. 3.2.5), peut-être avec des buts humoristiques pour les premiers et persuasifs pour les autres. Nous nous attendons aussi à des comparaisons de nature métaphorique, car celles-ci, en ajoutant du côté subjectif de l'auteur, permettent d'attirer l'attention du lecteur (Čechová, Krčmová & Minářová, 2008, p. 251).

Enfin, pour les **textes de spécialité**, nous envisageons une fréquence considérablement minoritaire de ce type de construction et si nous en trouvons quelques échantillons, nous nous doutons qu'il s'agira des locutions fortement figées. Certes, les comparaisons courantes vont dominer ici.

Pour pouvoir parler de la fréquence, nous allons nous servir des **corpus linguistiques**. Ce qui est l'atout majeur de la linguistique de corpus, c'est qu'elle représente aujourd'hui une méthode très efficace dans l'étude fréquentielle d'un lexème ou des séquences entières sur la base de l'usage et donc de la langue authentique (Cvrček & Richterová, 2014)⁴.

En premier lieu, nous allons travailler sur le corpus parallèle **InterCorp version 7** (French) et les corpus monolingues **L'Est républicain version 1**⁵ et **frWaC**⁶. Grâce aux corpus cités, nous allons créer des « **sous-corpus** » de travail, répartis d'après les types de textes que nous avons à disposition. Nous allons chercher la construction « *être* + ADJ +

⁴ [En ligne] [Consulté le 31.03.2016].

Disponible sur : <<http://wiki.korpus.cz/doku.php?id=pojmy:korpus&rev=1416829573>>

⁵ La version 2 date du 11.04.2016.

⁶ Tous les corpus consultés sont synchroniques.

comme » dans chacun des sous-corpus et, à partir des résultats trouvés, nous pourrions comparer **la fréquence** de la construction au niveau des types de textes particuliers.

En second lieu, nous allons vérifier les résultats reçus en ajoutant un autre corpus monolingue, à savoir *FRANTEXT*. Vu le caractère notamment littéraire de ce corpus, la recherche sera divisée encore en genres littéraires différents. Ainsi, nous pourrions voir également si le genre littéraire joue un rôle au niveau fréquentiel de la construction.

En troisième lieu, nous allons procéder à une « analyse manuelle » du **contexte droit** pour voir quelle suite se montrera comme la plus fréquente dans les différents sous-corpus. En effet, comme analysé dans le sous-chapitre 3.2.3, le contexte droit qui suit *comme* n'est pas uniforme du point de vue syntaxique. C'est pourquoi l'un de nos objectifs est de découvrir non seulement la fréquence de la construction dans les types de textes particuliers, mais aussi de chercher quelle est la suite la plus fréquente dans tel ou tel type de texte.

En dernier lieu, à la suite de l'étude du contexte droit, nous allons restreindre notre analyse à la **comparaison de haut degré** et grâce à des listes fréquentielles des adjectifs figurant dans la construction, nous allons essayer de trouver parmi nos résultats des locutions qui puissent être figées.

4.2 1^{ère} phase – « *être* + **ADJ** + *comme* + **X** »

Dans un premier temps nous allons considérer la construction comparative avec la suite « X » et donc avec toutes les suites possibles. Dans un second temps, l'analyse de la présente construction va nous conduire dans le sous-chapitre 4.3 jusque l'analyse d'une structure bornée, concrètement « *être* + **ADJ** + *comme* + **SN** » que Kokelberg (1993, p. 75) dénomme « le degré zéro de la comparaison ».

4.2.1 Fréquence de la construction selon le type de texte

Grâce à la base de données du *Corpus national tchèque*⁷, opérant sur l'outil de recherche « **KonText** », nous allons nous servir de trois corpus différents. Tout d'abord il nous faudra définir nos sous-corpus de travail d'après les types de textes à partir desquels les corpus sont composés.

Tout d'abord, il est nécessaire de préciser la composition textuelle des corpus consultés. En ce qui concerne *L'Est républicain* et *frWaC*, le premier ne regroupe que des

⁷ [En ligne] Disponible sur : < <http://www.korpus.cz> >

textes journalistiques et l'autre que des textes tirés du web. Vu l'homogénéité de ces corpus, il n'est pas nécessaire de diviser les textes.

Par contre, comme le corpus *InterCorp* contient des textes variés, nous avons créé les sous-corpus suivants :

I. faisant partie du « noyau » de corpus

- **belles-lettres** (sélection de « core – fiction »)⁸
- **textes de spécialité** (sélection de « core – nonfiction »)⁹

II. composant la « collection » du corpus¹⁰

- **sous-titres** (subtitles)
- **textes juridiques** (Acquis communautaire)¹¹
- **textes journalistiques** (PressEurop, Syndicate)¹²
- **Europarl** (textes provenant des sessions du Parlement Européen)¹³

De même, à part la composition textuelle des corpus qui est très importante, la **taille**, mesurée quant au nombre de positions¹⁴, joue un grand rôle aussi. Or, la grande taille d'un corpus n'implique pas forcément une grande **diversité lexicale**. En effet, nous pouvons obtenir les données de cette dernière en nous servant du rapport **types/tokens**, indiqué dans le tableau en-dessous :

Corpus	N° de types	N° de tokens	Taux de variabilité
L'Est Républicain	640 437	146 544 420	0,00437
InterCorp	430 211	98 885 271	0,00435
frWaC	5 182 079	1613814684	0,00321

Tableau 4.1 : Taux de variabilité lexicale des corpus *InterCorp*, *L'Est républicain* et *frWaC* reçu du rapport types/tokens

⁸ Il faut souligner que le sous-corpus des belles-lettres intègre surtout des romans et n'inclut pas les textes poétiques.

⁹ Nous pouvons y trouver par exemple des œuvres de Saussure, Bourdieu ou Foucault.

¹⁰ *InterCorp* contient des textes originaux ainsi que des traductions. L'alignement des textes du noyau est revu de manière semi-manuelle alors que la collection est alignée de façon automatique. Le noyau contient, en plus des belles-lettres et des textes de spécialité, des textes de vulgarisation scientifique pour les enfants et les bandes dessinées *Astérix* de Goscinny (cinq séries).

¹¹ Ensemble de textes juridiques utilisés en droit de l'Union Européenne

¹² Sites web journalistiques

¹³ Contient les sessions des années 2007 – 2011. Il s'agit des textes « pseudo oraux », car ceux-ci sont écrits préalablement à la session.

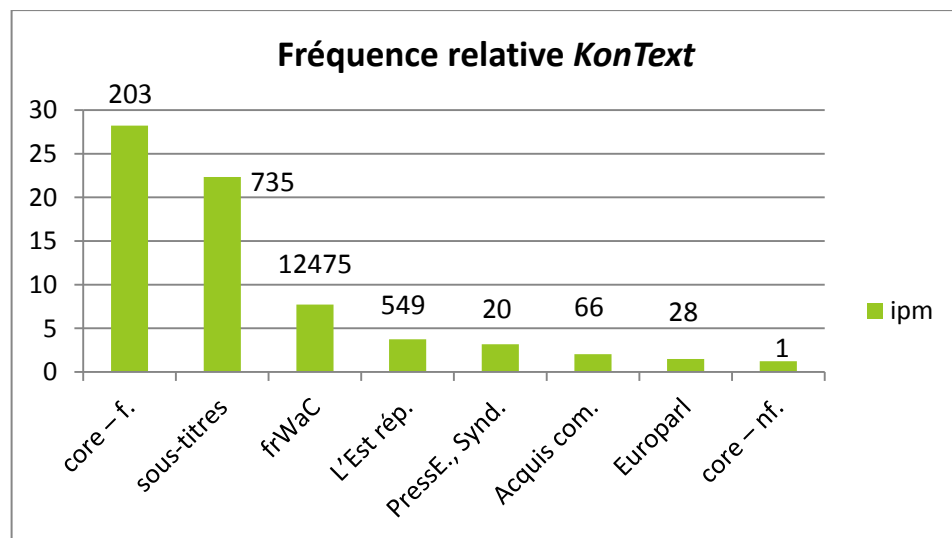
¹⁴ La position signifie le mot ainsi que la ponctuation.

Dans le tableau B.1 de l'**Annexe B** nous pouvons trouver également les données sur la taille de chaque sous-corpus créé.

Paradoxalement, le tableau 4.1 révèle que parmi les trois corpus, *frWaC* est justement le corpus qui reflète la plus grande taille, mais la plus petite variabilité lexicale. Par contre, c'est *L'Est républicain* qui l'emporte quant au lexique. Ceci renverse en partie notre hypothèse que les textes journalistiques englobent surtout des structures figées et stéréotypées (cf. 4.1). De plus, nous nous attendons à beaucoup de répétitions dans *frWaC*.

Ensuite, par le biais de l'outil de recherche *KonText*, nous avons cherché l'expression de séquence « [lemma="être"] [tag="ADJ"] [word="comme"] »¹⁵ dans chaque sous-corpus.

Observons la répartition fréquentielle de la construction dans le graphique suivant¹⁶ :



Graphique 4.1 : Fréquence relative (ipm) de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *KonText*, avec les fréquences absolues indiquées à l'intérieur du graphique

En effet, comme les sous-corpus consultés ne sont pas identiques quant au nombre de textes et de positions (cf. Tableau B.1), pour pouvoir les comparer il faut avoir recours plutôt à la fréquence relative qu'absolue¹⁷. La mesure « **ipm** » des sous-corpus particuliers nous indique les résultats plus fiables par rapport à leur taille, car la formule destinée à son calcul est la suivante :

¹⁵ Nous avons opté pour la langue de corpus de recherche (« **CQL** ») qui permet de faire la recherche combinée des **lemmes** et des **codes grammaticaux**.

Nous avisons qu'en nous limitant à cette expression de séquence nous nous privons des résultats que nous pourrions obtenir si nous ajoutions des positions libres : [lemma="être"] {0, n} [tag="ADJ"] [word="comme"]. En effet, ce sont par exemple des négations qui ne vont pas être prises en compte ainsi.

¹⁶ Les fréquences exactes absolues et relatives pour chaque sous-corpus figurent dans le tableau B.2 de l'**Annexe B**.

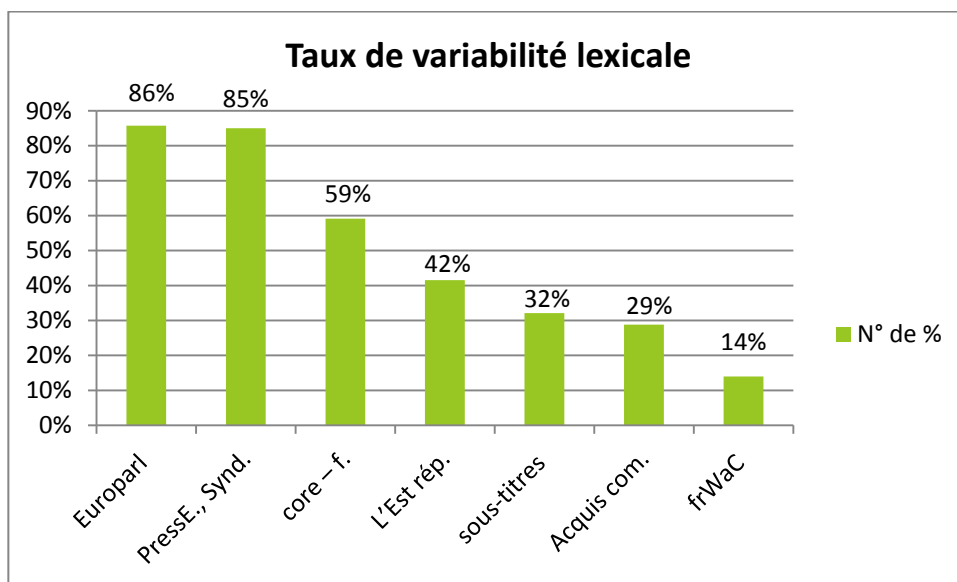
¹⁷ Désormais, nous appellerons les corpus *L'Est républicain* et *frWaC* également « sous-corpus ».

$$\text{ipm} = \frac{\text{n}^\circ \text{ d'occurrences}}{\text{n}^\circ \text{ de positions}} \times 1\,000\,000$$

Le graphique 4.1 nous renseigne que nous avons confirmé l'hypothèse fréquentielle de la construction, à savoir qu'elle est la plus fréquente dans les textes littéraires. Pour ce qui est des sous-titres, la seconde place pourrait résulter de l'usage quotidien des comparaisons courantes, du fait que nous considérons ce sous-corpus comme transcription de la langue orale et comme reflet du français parlé en quelque sorte. Sans compter que les scénarios dans la langue originale sont rédigés au préalable. Toutefois, les expressions clichées ou défigées ne sont pas exclues dans ces textes.

Une rupture considérable des deux premiers sous-corpus avec les autres par rapport à la fréquence relative est bien visible dans le graphique. Comme prévu, les textes sur Internet et les textes journalistiques se situent à peu près au milieu.

Les fréquences modestes des textes venant du domaine juridique et du *Parlement Européen* ne nous surprennent pas vraiment. Il s'agit des domaines où la clarté et la précision sont absolument requises. C'est la raison pour laquelle la comparaison de ressemblance introduite par *comme* ne devrait pas y occuper une place majeure. En effet, dans tous textes de spécialité en général, il n'est pas très convenable d'utiliser les mots polysémiques ou homonymiques (Čechová, Krčmová & Minářová, 2008, p. 219).



Graphique 4.2 : Taux de variabilité lexicale de la construction « être + ADJ + comme » en pour-cents dans les sous-corpus sélectionnés de *KonText* (excepté *core – nonfiction*), reçu du rapport types/tokens

Dans le graphique 4.2 nous introduisons les pourcentages exprimant le taux de variabilité lexicale à l'égard de la construction cherchée dans chaque sous-corpus¹⁸. Bien que les sous-corpus *Europarl* et *PressEurop & Syndicate* s'avèrent être les plus riches lexicalement, le nombre de résultats au total est très limité pour que nous puissions l'envisager comme relevant. Sinon, les textes littéraires confirment notre hypothèse quant à leur niveau d'originalité et du lexique peu stéréotypé.

Par contre, ce qui corrobore le taux de variabilité le plus élevé du corpus *L'Est républicain*, c'est justement le pourcentage de variabilité peu négligeable parmi les occurrences comparatives. Ceci peut découler du fait que d'un côté, le style journalistique suit une forme d'expression bien établie car les sujets se répètent souvent, mais de l'autre côté il tend intentionnellement à rompre avec des expressions automatisées (Čechová, Krčmová & Minářová, 2008, p. 249). De même, c'est un style très dynamique qui change continuellement dans le temps (op. cit., p. 246). Les locutions actualisées peuvent ainsi, devenir des clichés en peu de temps elles aussi.

Pour conclure, à ce que nous ne nous attendions pas, c'est que la construction est très peu fréquente dans tous les sous-corpus utilisés. Pour ce qui est des textes de spécialité, le corpus ne contient qu'un seul résultat issu de l'emploi illustratif de *comme* (cf. 1.4) qui est le suivant :

- (23) « [...] *qu'ils soient monosyllabiques* [les radicaux] **comme** *dans le sanskrit ou polysyllabiques comme l'hébreu* [...] »
(Foucault, M., *Les mots et les choses*, 1990)¹⁹

Bien évidemment, une seule occurrence ne nous dit rien sur le caractère lexical du domaine. C'est aussi pourquoi dans le sous-chapitre suivant, nous envisageons de consulter également un autre corpus plus riche en ce qui concerne les textes de spécialité.

Bref, les hypothèses que nous avons présentées ont été faites seulement sur de purs résultats fréquentiels. Ce qui compte encore c'est la nature des occurrences particulières, car sans un tri manuel, il peut y avoir beaucoup de « bruits ». Ces derniers vont être éliminés dans le sous-chapitre 4.2.3.

¹⁸ Dans le tableau B.2 de l'Annexe B nous pouvons trouver les nombres exacts de types et de positions pour chaque sous-corpus.

¹⁹ Pour cet exemple ainsi que pour tous les exemples suivants nous introduisons la date de publication qui peut être différente de la date d'origine. Dans *InterCorp*, il nous est arrivé parfois que la date d'origine y était manquante. Dans *FRANTEXT*, nous avons vérifié qu'il y figure la date de publication.

4.2.2 Fréquence de la construction selon le genre littéraire

Le nombre de résultats trouvés dans les corpus consultés dans 4.2.1 n'est pas suffisant pour pouvoir en déduire des conclusions. Il est donc préférable de consulter également un autre corpus monolingue, à savoir *FRANTEXT*. Ce qu'il faut mentionner, c'est que nous avons fait notre recherche seulement dans une partie du corpus, concrètement dans le *FRANTEXT catégorisé*, vu que celui-ci comprend des codes grammaticaux. Par opposition aux autres corpus utilisés (cf. Tableau 4.1), sa taille est égale à 127 515 681 positions²⁰.

Ce corpus regroupe des textes surtout à partir du XVI^{ème} siècle jusqu'au présent. Malgré cela, comme notre recherche n'est pas basée sur l'aspect diachronique, nous allons restreindre la datation des textes à partir du 20^{ème} siècle jusqu'aux textes les plus récents que le corpus nous offre, à savoir jusqu'à 1997²¹.

L'atout majeur de *FRANTEXT* est qu'il nous permet de diviser des genres littéraires particuliers. Aussi ont été sélectionnés le sous-corpus du **roman**, de la **poésie**, du **théâtre** et des textes de spécialité issus des **traités** et des **essais**²². Le nombre de romans étant à première vue très élevé, nous avons fractionné les résultats par périodes de vingt ans. Les autres sous-corpus n'ont pas été divisés.

Pour chercher la construction étudiée, nous utilisons l'expression de séquence suivante : « **&cêtre &e (g=A) comme** ».

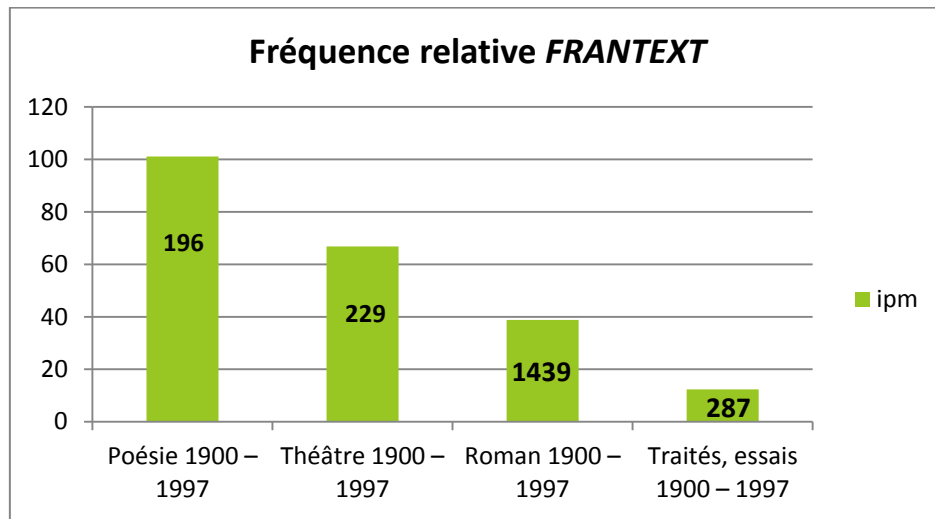
Afin de pouvoir comparer les fréquences des genres littéraires particuliers, nous allons nous servir de nouveau de la fréquence relative. Si nous observons le graphique 4.3, nous nous apercevons que le nombre d'occurrences n'est pas réparti de façon équilibrée. Les résultats trouvés nous confirment notre hypothèse quant au genre littéraire. En effet, il n'est pas étonnant que la poésie se sert de cette structure comparative le plus parmi tous les

²⁰ Seulement la partie catégorisée du corpus. Le *FRANTEXT intégral* comprend 285 923 119 mots (= positions). Le *FRANTEXT catégorisé* représente donc environ 45 % de l'ensemble du corpus.

²¹ Le *FRANTEXT intégral* qui contient aussi des textes sans codage grammatical comprend des textes plus récents.

²² Là-dessus, nous voulons faire un point sur la recherche incohérente dans *FRANTEXT*. En effet, non seulement la catégorisation des genres littéraires n'est pas clairement définie, mais aussi le catalogue des textes d'après la date n'est pas bien réglé. Dans notre cas, le nombre de textes était différent selon que nous ayons borné la datation des textes de 1900 – 1997 ou que nous ayons laissé la deuxième borne ouverte. Comme le corpus ne contient que les textes jusqu'à 1997, sans ou avec la borne finale, le nombre de résultats devrait être le même. Pourtant, lorsque nous avons introduit la borne finale, au lieu d'avoir 80 textes dans le sous-corpus des traités et essais, nous en avons eu 81. Le texte qui est apparu de plus était un texte de Baudelaire datant de 19^{ème} siècle.

sous-corpus étudiés du fait que c'est le genre opérant surtout sur le jeu de la langue et saturé de figures de style variées.



Graphique 4.3 : Fréquence relative (ipm) de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *FRANTEXT*, avec les fréquences absolues indiquées à l'intérieur du graphique

Le genre dramatique est à la deuxième place. La poésie le dépasse pourtant de presque moitié. Si nous généralisons nos connaissances sur ce genre, ce dernier est réputé par son emphase voire hyperbole. Dès lors il se peut que des locutions métaphoriques y soient présentes dont un grand nombre de locutions figées voire rabâchées.

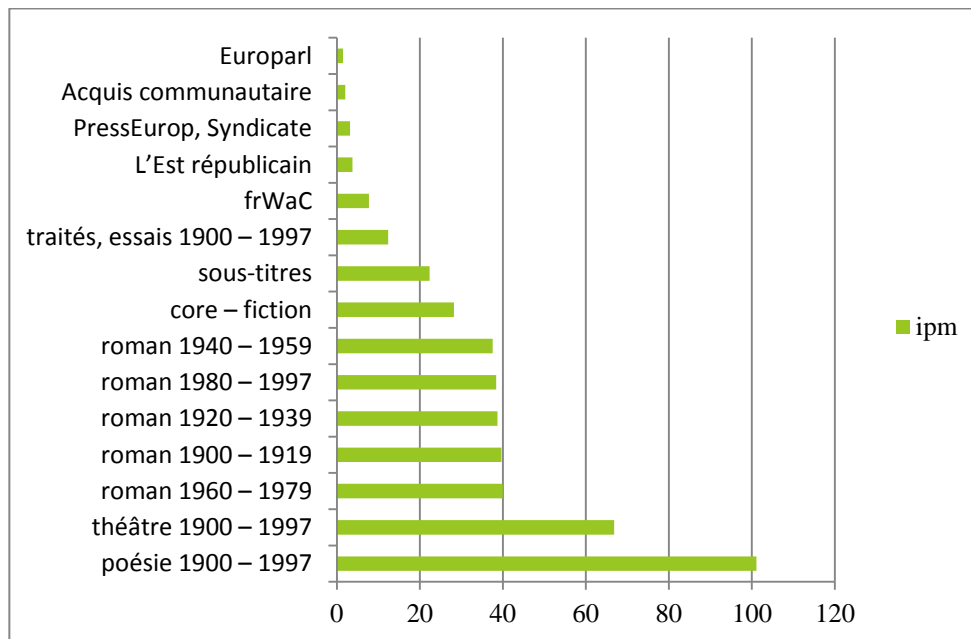
En ce qui concerne le genre romanesque, nous notons que la fréquence est encore plus petite que pour les deux premiers genres. Les fréquences relatives des sous-corpus de différentes périodes sont à peu près similaires (cf. Tableau B.3), en dépit du fait que le nombre de romans de différentes périodes n'est pas identique (cf. Tableau B.4) et que le sous-corpus des années 1920 – 1939 surpasse de moitié toutes les autres périodes par rapport à la fréquence absolue²³. Somme toute, nous supposons que la mesure de l'originalité dépend ici en bonne partie du style individuel, propre à un auteur.

Finalement, la fréquence minimale de la construction dans les textes de spécialité corrobore les résultats trouvés déjà dans le corpus *InterCorp* (cf. Graphique 4.1 dans 4.2.1), sauf qu'ici la nature des textes est un peu différente. En effet, nous pouvons y trouver des textes très hétérogènes, des textes traitant de la météorologie jusqu'à des textes sur l'art. Parmi ces derniers, il y figure par exemple *L'Art poétique* de Claudel, les manifestes de Tzara ou le théâtre d'Artaud. Il se peut donc que ce sous-corpus contienne

²³ Ceci s'explique par le fait que la période de l'entre-deux-guerres apporte, entre autres, une production littéraire très féconde, notamment en ce qui concerne le genre romanesque.

aussi des comparaisons métaphoriques. Néanmoins, nous en restons toujours sur le niveau hypothétique, car seulement l'analyse qualitative des résultats dans les sous-chapitres suivants peut nous en dévoiler plus.

Dans l'intention de mieux voir quels sous-corpus de *KonText* ainsi que de *FRANTEXT* sont ceux qui comprennent le plus d'occurrences issues de la construction cherchée, nous avons inséré la fréquence relative de chaque sous-corpus un graphique :



Graphique 4.4 : Comparaison des fréquences relatives (ipm) de la construction « être + ADJ + comme » dans tous les sous-corpus sélectionnés de *KonText* et de *FRANTEXT* (excepté *core – nonfiction*)

Du graphique 4.4 il est clair que dans *FRANTEXT* la construction est beaucoup plus fréquente que dans les trois autres corpus. Ceci découle du fait que celui-là est formé en majeure partie par des textes littéraires. Seulement les sous-corpus des belles-lettres et des sous-titres dans *InterCorp* précèdent le sous-corpus des traités et essais dans *FRANTEXT*.

Or, comme mentionné déjà avant, la taille des sous-corpus est très variée si bien que les différences fréquentielles de la construction sont à relativiser. De plus, ce qui peut entrer encore en jeu c'est l'influence des **idiolectes** de certains auteurs qui utilisent un mot ou une construction plus souvent qu'il est standard.

En définitive, comme nous allons le voir dans le sous-chapitre suivant, toutes les constructions trouvées avec l'expression de séquence donnée ne sont pas nécessairement des structures comparatives.

4.2.3 Résultats « intrus »

Dans l'intention de faire une analyse qualitative de nos résultats, nous avons créé des échantillons aléatoires au nombre de 250 occurrences pour chacun des sous-corpus dans lesquels la fréquence était supérieure à 250 occurrences, à savoir dans *frWaC*, *L'Est républicain* et les sous-titres dans *InterCorp*. De même pour les occurrences de *FRANTEXT*, les résultats analysés dans les sous-corpus du roman 1920 – 1939, 1940 – 1959, 1960 – 1979 et des traités et essais ont été limités au nombre 250 chacun²⁴. Les autres sous-corpus ont été analysés dans leur ensemble (cf. Tableaux B.2 et B.3).

Or, il faut souligner qu'au cours de notre recherche il fallait éliminer des résultats « intrus » qui ne correspondaient pas à notre objectif. En moyenne, les **bruits** constituaient 18 % des résultats de tous les sous-corpus. Pour pouvoir nous en débarrasser, nous avons dû passer par toutes les lignes de concordance.

➤ Tout d'abord, nous avons supprimé des **lignes de concordance doublées**, ce qui était le cas des corpus *L'Est républicain* (6 occ. sur 250) et *frWaC* (1 occ. sur 250)²⁵. Cela peut être dû à l'absence de la correction manuelle des occurrences, ses dernières étant téléchargées automatiquement du web.

➤ Puis, il s'agissait **des adjectifs** utilisés en **fonction adverbiale** plusieurs fois dans la suite « *c'est juste comme ça* » (sous-titres) qui veut dire « c'est exactement comme ça » ou des adjectifs faussement liés non pas au verbe *être*, mais à sa variante **nominale** *l'être*, par exemple « *Le développement de l'être humain comme individu [...]* » (*frWaC*)²⁶. Cela montre l'existence des erreurs dans le **codage grammatical**, présentes dans tous les corpus étudiés.

²⁴ Il existe une méthode statistique selon laquelle il est conseillé de faire trois échantillons différents qui compteraient ensemble 10 % du total des occurrences pour que les résultats soient fiables (si le nombre de résultats est inférieur à 10 000). Dans ce travail nous avons opté pour un échantillon de 250 résultats qui nous paraissait suffisant en nous rendant compte des constructions répétitives ainsi que de la fréquence peu élevée des autres sous-corpus.

²⁵ Nous introduisons, plutôt que des pourcentages, toujours des fréquences absolues pour les nombres d'occurrences peu importants.

²⁶ Aux exemples cités nous indiquons toujours leur source. Nous introduisons le sous-corpus à partir duquel nous les avons tirés. Ainsi, s'il est indiqué « théâtre », « poésie » ou « roman », nous faisons référence à *FRANTEXT*. Par les « belles-lettres » nous pensons à *InterCorp*. Les métadonnées que nous ajoutons sont l'auteur, le titre de l'œuvre et l'année de publication pour les textes littéraires et la date de publication pour les textes journalistiques. Quant aux sous-titres, nous y ajoutons le titre du film et l'année de parution.

➤ De plus, le vers libre ainsi que l'absence de ponctuation sont très courants dans le genre poétique. C'est pourquoi il peut arriver parfois que l'outil de recherche interprète faussement **l'ordre des mots** dans une phrase.

(24) « *Tant l'énervement était général*

Comme de simples allumettes

Les gens sortaient de chez eux [...] » (poésie, Aragon, L., *Les destinées de la poésie*, 1926)

En effet, dans la séquence ci-dessus, la suite après *comme* ne se lie pas à *être général*, mais à *Les gens* dans le vers suivant.

➤ Par ailleurs, pour ce qui est des configurations **non-comparatives**, elles étaient assez nombreuses. En effet, les **participes** identifiés comme adjectifs étaient courants. Il s'agissait des participes dont le paradigme est nécessairement construit avec une préposition. Pour cette raison, dans les **textes juridiques**, aucune des occurrences ne correspondaient à la construction cherchée. *Comme* figurait comme préposition dans des phrases avec le participe présent :

(25) « [...] *la valeur normale de chaque type de produit a été déterminée comme étant la moyenne pondérée de tous les prix de vente [...] »* (Acquis),

avec le participe passé : « *est apparu comme pour* » (FRANTEXT) ou avec les participes finissant en suffixe *-able*, par exemple « [...] *sont applicables comme condition obligatoire [...] »* (Acquis).

➤ En outre, des suites verbales étaient très souvent employées avec des adjectifs tels que *fou*, *drôle*, *étonnant* et bien d'autres qui expriment l'étonnement ou une constatation subjective au niveau des **phrases exclamatives**²⁷.

(26) « *C'est fou comme le temps passe vite [...] »* (L'Est républicain, 22.02.2003)

(27) « *C'est drôle comme on oublie vite un nom. »* (sous-titres, *Domestic Disturbance*, 2001)

Les intrus de ce type ont été présents surtout dans les textes qui reflètent le discours subjectif avec beaucoup d'exclamations, tels que les sous-titres (12 %) ou les textes sur Internet (8 %), mais aussi dans les textes journalistiques de *L'Est républicain* (9 %). Il est à noter que la plupart de ces phrases figuraient sans point d'exclamation (cf. 3.1).

²⁷ Ceci confirme l'affirmation que l'exclamation est un des emplois principaux de *comme* (cf. Chapitre 1).

Même si nous éliminons ce type de séquences de notre recherche, il existe des liens entre elles et la comparaison. En effet, ce qui est partagé comme trait, c'est la sémantique de haut degré (Fuchs, 2014, p. 161).

➤ Finalement, des structures qui se sont répétées plusieurs fois étaient des constructions aux prédicats attributifs avec le sens d'**appartenance à une classe**. Regardons un exemple :

(28) « *C'est beau comme prénom* [sic] *tu ne trouves pas ?* » (belles-lettres, Gavalda, A., *Ensemble, c'est tout*, 2006)

En effet, dans cet exemple le prénom est pris comme le représentant de la classe des prénoms (Portine, 1995, p. 373). En général, l'élément qui suit *comme* est sans article (op. cit., p. 378)²⁸. La majeure partie de ce type de phrases ont été utilisés dans le contexte du français parlé au niveau des textes tirés du web (13 %), des sous-titres (10 %) et quelques échantillons ont été trouvés aussi dans les belles-lettres (5 %), *L'Est républicain* (7 %) et *Europarl* (11 %).

Il est probable qu'en littérature cela reflète le discours direct d'un personnage parlant de sorte que son langage lui soit typifiant par rapport à son rôle dans le texte²⁹.

Les substantifs exprimant l'appartenance à une classe de métier, de religion, de parti politique et d'autres de ce type ont été faussement codés comme adjectifs, probablement parce que ces substantifs se construisent sans article dans ce type de phrases, par exemple « [...] *tout le monde serait ouvrier comme autrefois* [...] » (belles-lettres, Ouředník, P., *Europeana, Une brève histoire du XXe siècle*, 2004).

Le tableau 4.2 montre le nombre de résultats analysés dans le sous-chapitre suivant après la suppression des occurrences « intruses » non-comparatives. Etant donné la fréquence (un seul résultat) du sous-corpus des textes de spécialité dans *InterCorp* (cf. 4.2.1) et la totalité des occurrences « bruisantes » de l'*Acquis communautaire*, ces deux sous-corpus vont être exclus de nos prochaines analyses. Beaucoup de bruits sont observables aussi dans les sous-corpus *Europarl*, *PressEurop & Syndicate*, *frWaC* et dans les traités et essais. Dans les deux premiers sous-corpus c'est à cause de petites fréquences

²⁸ Portine (1995) consacre l'article entier à la distinction entre la comparaison et l'appartenance à une classe.

²⁹ **Exemples** : « *Hé ben, vous me croirez pas, mais la classe entière s'est mise à renauder que c'était dégueulasse comme sujet infaisable* [,] *un truc exprès pour les coincer* [...] » (roman, Seguin, F., *L'Arme à gauche*, 1990)

« *Croyez-vous qu'il aurait encore assez de patte pour attraper ça? Est-ce assez fort! Et puis, c'est beau comme matière* [,] *ça serait amusant à tripoter.* » (roman, Proust, M., *Sodome et Gomorrhe*, 1922)

au total, ces dernières défigurant les pourcentages, et ce, à cause de la nature spécifique de leurs textes. Les bruits dans *frWaC* résultent principalement de l’usage fréquent des structures exclamatives et attributives.

sous-corpus <i>FRANTEXT</i>	f absolue	N° de %	sous-corpus <i>KonText</i>	f absolue	N° de %
roman 1900 – 1919	187	92 %	core – fiction	162	80 %
roman 1920 – 1939	218	87 %	sous-titres	178	71 %
roman 1940 – 1959	215	86 %	frWaC	167	67 %
roman 1960 – 1979	196	78 %	L’Est républicain	181	72 %
roman 1980 – 1997	200	87 %	PressEurop, Syndicate	12	60 %
poésie 1900 – 1997	181	92 %	Acquis communautaire	0	
théâtre 1900 – 1997	206	90 %	Europarl	16	57 %
traités, essais 1900 – 1997	170	68 %	core – nonfiction	0	

Tableau 4.2 : Fréquences absolues et pourcentages des occurrences analysées de la construction « être + ADJ + comme » dans tous les sous-corpus sélectionnés de *KonText* et de *FRANTEXT* après la suppression des « bruits »

Somme toute, il faut souligner que le codage grammatical n’est pas fiable à cent pour cent. C’est pourquoi il est nécessaire de faire toujours une analyse manuelle des résultats trouvés afin de ne pas obtenir de résultats trompeurs. Dans la partie à venir nous allons mentionner toutes les autres suites qui peuvent être placées après *comme* et cette fois-ci il s’agira des structures purement comparatives.

4.2.4 Contexte droit

Comme tous les résultats ne nous convenaient pas, nous avons dû en éliminer une bonne part de tous les sous-corpus (cf. 4.2.3). Ainsi, les textes juridiques et les textes de spécialité dans *InterCorp* ne vont pas être considérés faute de structures comparatives dans les premiers et faute de résultats suffisants pour les autres.

Ce qu’il faut remarquer à propos des textes dans l’*Acquis communautaire*, c’est que la suite prédominante de *comme* est la structure participiale (cf. (25)), mais celle-ci est sans valeur approximative dont nous avons parlé dans 3.2.3. Dans les autres sous-corpus, ce type de construction est, en revanche, très peu fréquent. La raison en est le style très spécifique employé dans les textes juridiques, caractérisé par des structures lourdes, complexes et difficilement compréhensible aux « non-initiés ». Les textes de spécialité de *FRANTEXT* ne démontrent pas non plus cette tendance.

Après avoir observé nos résultats des corpus de plus près, nous pouvons constater que tous les contextes syntaxiques droits mentionnés dans 3.2.3 ont été trouvés (SV, SN,

complément circonstanciel, adjectif/participe, combinaison de constituants). Comme les suites des corpus étaient plus variées avec des nuances sémantiques, nous avons élargi le tableau classifiant³⁰ :

sous-corpus	SN	SN étendu	SV	pron.	ça/cela	tout	PP/ADJ	ccir.	approx.	add.	exem.
roman	780	187	83	89	29	23	0	57	3	6	3
poésie	157	50	14	6	0	4	0	8	0	0	0
théâtre	170	25	16	10	6	4	0	3	0	0	0
traités, essais	98	10	35	28	1	2	2	13	0	1	3
core – f.	116	23	25	10	6	5	0	8	0	0	0
sous-titres	131	8	2	15	19	10	1	2	0	0	0
frWaC	98	13	28	7	9	10	3	8	0	5	7
L'Est rép.	114	31	38	6	6	4	0	23	0	5	8
PressE., Synd.	8	0	2	0	1	0	0	1	0	0	1
Europal	8	7	4	1	2	0	0	2	0	0	1

Tableau 4.3 : Répartition des occurrences analysées d'après le contexte droit de *comme* dans la construction « être + ADJ + *comme* » (après la suppression des bruits) dans les sous-corpus sélectionnés de *KonText* (excepté l'*Acquis communautaire* et *core – nonfiction*) et de *FRANTEXT*

➤ La suite X la plus fréquente au sein de tous les sous-corpus est décidément le **syntagme nominal**³¹, en poésie représentant 83 % et en genre dramatique 81 % de toutes les suites. Parmi les sous-corpus d'*InterCorp*, *L'Est républicain* et *frWaC*, les sous-titres dépassent les autres sous-corpus avec 73 % des structures nominales.

D'ailleurs, les SN figurent **étendus** ou sans expansions. Pour ce qui se rapporte aux expansions, nous avons observé parmi nos résultats les SN étendus par des épithètes³², des compléments du nom³³ et des propositions relatives³⁴ aussi. Même s'il ne s'agit pas des expansions en tant que telles, nous avons inséré dans ce groupe également des SN combinés avec des circonstants³⁵. Les SN étendus ensemble avec des suites combinées ne comptent que 21 % de toutes les occurrences de tous les corpus.

³⁰ Dans ce cas, nous avons séparé la suite pronominale de la suite nominale et divisé des structures pronominales spécifiques *comme ça/cela* et *comme tout*.

³¹ « [...] *son cœur est content comme une sauterelle* » (poésie, Noailles, A. de, *Le cœur innombrable*, 1901)

³² **Exemple** : « *Comment faire pour être heureux comme un petit enfant candide* » (poésie, Apollinaire, G., *Alcools*, 1913)

³³ « [...] *leur visage était dur comme les pierres de la vallée* » (roman, Le Clézio, J.M.G., *Désert*, 1980)

³⁴ « *Rachaella parlait avec volubilité et était joyeuse comme une enfant qui aurait découvert son cadeau de Noël* » (frWaC)

³⁵ Dans 3.2.3 nous avons traité la combinaison des constituants individuellement.

Exemple : « *Ils sont purs comme le ciel au printemps*. » (théâtre, Lenormand, H-R., *Le Simoun*, 1921)

Sinon, le reste des SN sont sans expansions et il s'agit des **structures elliptiques**, présentées dans 3.2.3. Ce qui est remarquable, c'est que les structures elliptiques les plus nombreuses figurent dans les sous-titres (94 %), les traités et essais (90 %), *frWaC* (87 %) et le théâtre (85 %). Ceci indique que notre hypothèse sur la présence des comparaisons majoritairement figées pourrait être vraie, car plus les suites sont étendues, moins elles sont figées en principe.

Par contre, des SN étendus sont les plus nombreux dans les textes de poésie (32 %). Comme nous le verrons dans les sous-chapitres suivants et comme prévu au début de ce chapitre, la poésie est la source des comparaisons surtout originales et non-figées.

➤ La suite la plus fréquente après le syntagme nominal est le **syntagme verbal**, autrement dit une suite propositionnelle³⁶. D'abord, nous avons repéré des verbes repris de la proposition principale ainsi que les verbes supports (cf. 3.2.3.). Puis, des verbes différents de ceux de la principale y figurent aussi³⁷.

La comparaison métalinguistique n'y manque non plus. En plus des variantes du verbe *dire*, nous avons détecté des verbes tels que *témoigner*, *attester* ou *montrer*³⁸.

De plus, nous avons inséré sous cette catégorie également des propositions hypothétiques avec **comme si**³⁹.

(29) « *Ses jambes ne danseront plus, elles seront lourdes **comme si** ses chaussures de plage pesaient une tonne* » (roman, Belloc, D., *Kepas*, 1989).

En effet, cet exemple nous semble être sémantiquement proche du sens approximatif (cf. 1.4). Celui-ci est traité séparément dans les lignes suivantes. Portine (1996, p. 376) explique qu'avec *comme si* le rapport comparatif est exprimé par une hypothèse qui sert de référence à l'évaluation d'un jugement.

³⁶ Les suites infinitives (infinitif présent ou passé) sont incluses aussi, même si elles présentent le sujet de la phrase, par ex. « *Rien n'est difficile **comme de vivre en public** [...]* » (traités et essais, Guehenno, J., *Jean-Jacques T.2*, 1950) ou « [...] *mes armées étaient lasses **comme d'avoir porté un lourd fardeau***. » (Saint-Exupéry, A. De, *Citadelle*, 1944).

³⁷ **Exemple** : « *Tartuffe **est antipathique comme le souhaitait** Molière* » (L'Est républicain, 10.02.2003)

³⁸ **Exemple** : « *Son sommeil fut calme **comme en témoigne** la jolie enveloppe de coton demeurée aussi nette [...]* » (roman, Rheims, M., *Les greniers de sienne*, 1987)

³⁹ Ce qui est intéressant sur les propositions introduites par *comme si*, c'est qu'à la différence de *comme* temporel traité dans 2.1, elles sont compatibles avec tous les temps de l'indicatif ainsi que du subjonctif. Les plus courants en sont l'imparfait ou le plus-que-parfait et dans la langue littéraire aussi le subjonctif plus-que-parfait (Grevisse & Goosse, 2008, p. 1517).

➤ La **suite pronominale** (pron. dans le tableau 4.3) est aussi assez fréquente. Elle est la plus fréquente dans les traités et essais avec 29 %. Le plus souvent il est question des pronoms personnels seuls⁴⁰ ou des pronoms démonstratifs avec un complément :

(30) « *Les mœurs et leurs goûts étaient simples comme **ceux des villageois** [...]* »
(traités et essais, Dévigne, R., *Le légendaire de France*, 1942)

Les suites pronominales ont le même fonctionnement que les suites nominales. Plus précisément, les pronoms apparaissent étendus par des compléments du nom (cf. (30)) ou par des propositions entières⁴¹. La seule différence qui s'impose comme pertinente du point de vue sémantique, c'est que par opposition aux SN, les suites pronominales ne se prêtent pas à former des comparaisons figuratives (cf. 4.2.5).

Comme peut être suivi encore de *cela* ou de sa variante *ça* propre au style familier. En effet, ce type de structure est à l'oral naturellement accompagné par un geste (Rey & Chantreau, 1991), ce qui explique sa fréquence plus importante dans les sous-titres (19 occ.).

Le pronom indéfini *tout* figure comme la suite aussi et pareil comme pour *cela* et *ça*, le plus de ses occurrences sont enregistrées dans les sous-titres (10 occ.) et les textes du web (10 occ.). Dans le sous-chapitre suivant, *tout* sera présenté aussi du point de vue sémantique.

➤ Les comparaisons exprimant la ressemblance entre situations (cf. 3.2.3) ne sont pas non plus négligeables. Nous avons repéré tous les types de **circonstants** placés après *comme*, c'est-à-dire des compléments circonstanciels de lieu, de temps ou de manière. La majeure partie des occurrences vient de *L'Est républicain* (23 occ.), ce qui peut être connecté à l'usage commun des actualisateurs de discours comme les déictiques spatio-temporels dans le domaine de la presse.

➤ En outre, peu d'occurrences ont été trouvées aussi pour *comme* exprimant l'**addition** (add. dans le tableau 4.3), l'**exemplification** (exem. dans le tableau 4.3) ou l'**approximation** (approx. dans le tableau 4.3) (cf. 1.4). Premièrement, pour Portine (1995, p. 391), l'exemplification se situe dans cette construction à mi-chemin entre la

⁴⁰ **Exemple** : « *Elle est seule comme moi.* » (Sartre, J-P., *La Nausée*, 1938)

⁴¹ **Exemple** : « *Mais cette impression est trompeuse comme celle que donne un film de cinéma accéléré.* » (traités et essais, Ruyer, R., *Esquisse d'une philosophie de la structure*, 1930)

comparaison et l'appartenance à une classe⁴². Deuxièmement, concernant l'emploi coordonnant de *comme*, Fuchs (2014, p. 163) le considère aussi déjà à la limite du schéma comparatif. Bien entendu, tout dépend de la phrase concrète⁴³. Troisièmement, l'usage approximatif, qui repose sur le fait de qualifier une entité par un mot similaire faute de trouver le mot exact plus convenant (op. cit., p. 162), a été exprimé dans nos occurrences seulement par des suites adjectivales⁴⁴.

➤ De plus, l'**adjectif** ou le **participe** placés après *comme* ne supposent pas forcément une approximation (cf. 3.2.3). En effet, comme nous l'avons signalé au début de ce sous-chapitre, cette suite est le plus souvent rencontrée dans les textes juridiques, ces derniers éliminés de notre étude dans 4.2.3. Sinon, dans les autres types de textes, la suite est très rare (cf. Tableau 4.3)⁴⁵.

➤ Enfin, deux fois il est arrivé que *comme* était suivi de **point de suspension**. Le comparant reste donc, à notre connaissance, inexprimé. Or, il n'est pas exclu que dans le texte qui suit le comparant soit dévoilé. Ce qui est intéressant sur ce point est qu'il s'agit exactement de la même occurrence, à savoir « *Il dit : " Vous êtes beau comme ... " Logre sourit.* », trouvée une fois dans les belles-lettres d'*InterCorp* et l'autre fois dans le sous-corpus du roman des années 1960 – 1979 dans *FRANTEXT* chez Michel Tournier dans son œuvre *Le coq de bruyère* (1991 et 1978).

Dans le sous-chapitre suivant nous allons poursuivre notre étude sur des constructions comparatives avec *comme* en laissant de côté leur contexte syntaxique pour pouvoir chercher des sens figurés.

⁴² Il est à distinguer d'un côté des phrases comme « *Quelques artistes étaient présents comme Bau, guitariste de Cesaria Evora, ou Malaquias [...]* (L'Est républicain, 09.09.2002), où il est question d'une pure exemplification, de l'autre côté des phrases telles que « *Certains motifs seraient originaux comme les papillons, les rinceaux polychromes et le damier d'or [...]* (traités et essais, Fontaine, G., *La céramique française*, 1965) dans lesquelles la sémantique d'appartenance à une classe est plus marquante. En effet, on nomme ici les motifs qui font partie de la classe des motifs originaux.

⁴³ « *Dans l'idée [sic] ça peut être nul comme génial.* » (frWaC) vs. « *Les paysages sont splendides comme les acteurs d'ailleurs.* » (frWaC). Vu que le premier exemple manque de comparant, on ne voit pas le rapport de comparaison. Le deuxième, par contre, peut être reconstitué en « *Les paysages sont splendides comme le sont les acteurs* ».

⁴⁴ **Exemple** : « *Ils sont beaux comme neufs* » (roman, Dorgelès, R., *Les croix de bois*, 1919).

Fuchs (2014, p. 161) présente encore le schéma approximatif *être* + *comme* + SN, vu déjà dans 1.4 (cf. (3)), dans lequel il s'agit d'une dénomination de l'entité comparée, par ex. « *Il y a comme un défaut* ».

⁴⁵ **Exemple** : « [...] *ce texte était inconstitutionnel comme autorisant l'empiètement du pouvoir législatif sur le pouvoir judiciaire [...]* » (traités et essais, Vedel, G., *Manuel élémentaire du droit constitutionnel*, 1949). Le PP exprime ici la cause (cf. 2.2). Nous pourrions dire aussi « *Comme il autorisait l'empiètement du pouvoir législatif sur le pouvoir judiciaire, ce texte était inconstitutionnel.* ». Cet exemple fait cependant référence au domaine juridique.

4.2.5 Le haut degré

Nous avons vu tout à l'heure tous les X, c'est-à-dire tous les contextes droits qui figurent dans nos sous-corpus dans la construction « *être* + ADJ + *comme* + X ». Pourtant, notre intérêt n'est que d'examiner des structures comparatives figuratives (cf. 3.2.4). C'est la raison pour laquelle il nous faut d'ores et déjà, en reprenant la terminologie appropriée par Moline (2006, p. 17), séparer les « **fausses comparaisons** » des « **vraies comparaisons** », en d'autres termes les comparants à **valeur générique intensive** des comparants issus d'un référent présent dans une situation sans valeur intensive⁴⁶ :

(31) « [...] *son fils Jean-Marie était châtain comme son père.* » (roman, Bourget, P., *La Geole*, 1923)

vs.

(32) « *Elle est sourde comme un pot.* » (sous-titres, *The Curious Case of Benjamin Button*, 2008)

En effet, le premier exemple manque de valeur intensive, ce qui est contraint par l'usage des déictiques personnels et démonstratifs⁴⁷. Par contre, le deuxième exprime un haut degré de surdité en utilisant un comparant à parangon figé.

En suivant cette dichotomie sémantique, nous avons effectué un tri de tous les résultats selon notre jugement personnel. Dans l'ensemble il faut dire que ce n'était pas toujours facile de distinguer les deux types de comparaison :

(33) « *Le pacte était clair comme les yeux de Clémence* [...] » (roman, Genevoix, M., *Rroû*, 1931)

Le comparant est dans cet exemple un référent concret. Or, non seulement l'expression est métaphorique en mettant côte à côte une idée abstraite avec une partie du corps (cf. 4.3.1), mais aussi on appuie sur le haut degré de la clarté du pacte en sous-entendant que les yeux (bleu clair ?) de Clémence sont très clairs. Rien ne nous interdit non plus de le concevoir d'une manière ironique au cas où les yeux de Clémence seraient marron foncé. En effet, même un référent présent dans la situation d'énonciation peut devenir parangon pertinent au moment de l'énonciation (Moline, 2006, p. 18)⁴⁸.

⁴⁶ Autre terminologie introduite par Moline (2006, p. 17) : constructions figées/libres

⁴⁷ Une exception en peut être la locution « *être bête comme ses pieds* » à valeur générique malgré la présence du déictique possessif.

⁴⁸ Un autre exemple en est « [...] *la vie était rose comme sa chemise* » (roman, Lanzmann, J., *Le Têtard*, 1976). Cela ne veut pas dire qu'on veuille comparer la vie à la chemise, mais qu'on accentue la beauté de la vie en se servant de la couleur rose de la chemise.

Ce qui nous facilite la tâche, c'est que le syntagme nominal, suite la plus fréquente (cf. 4.2.4), nous fournit une forte majorité des occurrences de comparaison générique. Le reste est complété par quelques rares SV, par exemple « [...] *ta main est fine comme le soir est impalpable* » (poésie, Tzara, T., *Indicateur des chemins de cœur*, 1928), des pronoms étendus par des compléments, par exemple « *Sa voix était douce comme celle d'un ange.* » (sous-titres, *Ladyhawke*, 1985) et par des compléments circonstanciels « *Le silence est profond comme dans un carrefour de forêt* » (poésie, Claudel, P., *Connaissance de l'Est*, 1907).

KonText	f absolue	N° de % figuré	FRANTEXT	f absolue	N° de % figuré
sous-titres	110	70	poésie 1900 – 1997	131	78
core – fiction	96	63	théâtre 1900 – 1997	134	76
L'Est républicain	69	45	roman 1940 – 1959	131	66
frWaC	54	43	roman 1920 – 1939	117	59
Europarl	6	39	roman 1960 – 1979	116	63
PressEurop, Syndicate	5	38	roman 1980 – 1997	107	60
			roman 1900 – 1919	100	59
			traités, essais	43	32

Tableau 4.4 : Fréquences absolues et pourcentages de comparaison figurative à SN dans la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *FRANTEXT* et de *KonText* (excepté l'Acquis communautaire et core – nonfiction)

Le tableau 4.4 montre le nombre d'occurrences issues du SN à sens figuratif qu'il nous reste après l'élimination des comparants spécifiques et des suites autres que SN. A la suite d'une analyse de l'aspect sémantique des occurrences, nous en sommes venus à la distribution suivante. La comparaison de haut degré est supérieure à 59 %, excepté les traités et essais (32 %), dans tous les sous-corpus de *FRANTEXT*, la poésie (78 %) avec le théâtre (76 %) étant sur les deux premières places. Il est donc clair que dans les textes littéraires, la comparaison dominante sera la comparaison figurative.

Malgré notre réflexion sur la nature poétique possible des textes issus des traités et des essais, il faut avouer que l'hypothèse première que ces textes contiennent surtout des comparaisons courantes a été confirmée. En effet, la plupart des occurrences y sont liées à une situation concrète sans valeur intensive voire métaphorique.

- (34) « *Les théories politiques sont changeantes comme les circonstances qui les déterminent.* » (traités et essais, Bainville, J., *Histoire de France*, 1924)

Néanmoins, si nous y avons repéré quelques comparaisons figuratives, il s'agissait la plupart du temps de comparaisons assez originales comme « [...] *le feu du cœur est rouge comme du cinabre* [...] » (traités et essais, Caron-Hutin, *Les alchimistes*, 1959). Cela dépend, bien évidemment, du domaine duquel les textes viennent et de l'auteur concret. Une comparaison qui peut paraître expressive dans les textes de spécialité, peut s'avérer peu intéressante en littérature.

Pour ce qui est des autres corpus, la comparaison intensive domine dans les sous-corpus des sous-titres (70 %) et des belles-lettres (63 %). La fréquence dans les sous-titres va de pair avec la fréquence des phrases exclamatives, courantes dans ce sous-corpus (cf. 4.2.3). Le plus petit pourcentage est observable dans les textes journalistiques de *PressEurop* et *Syndicate* où seulement cinq résultats avec un sens figuré (38 %) ont été trouvés. Bien entendu, cela peut être dû à une faible quantité des occurrences au total. C'est pareil pour le sous-corpus *Europarl* avec six résultats figuratifs à parangon (39 %).

Nous avons, de même, détecté des tournures spécifiques telles que *comme tout/personne/pas un* qui marquent le plus haut degré ou une généralité absolue et pour lesquelles Fuchs (2014, p. 81) utilise le terme de « **pseudo-incomparabilité** ».

(35) « [...] *le vieux Kafka lui aussi était nerveux comme pas un* [...] » (belles-lettres, Hrabal, B., *La chevelure sacrifiée*, 1987).

Ce procédé a été trouvé aussi dans l'ordre inversé, comme par exemple dans « *rien n'est facile comme la vie dans ma tête jetée sur la plage fine* [...] » (poésie, Tzara, T., *Où boivent les loups*, 1932), où l'entité à comparer est placée dans la position du comparant et la phrase devient négative.

En ce qui concerne le schéma « X + comme + **tout** » (cf. 4.2.4), il exprime aussi le degré extrême, notamment avec les sens mélioratif (Rey & Chantreau, 1991, p. 231). Ce schéma apparaît le plus souvent dans les sous-titres (6 %) et les textes du web (6 %). Les adjectifs les plus utilisés dans cette construction sont les adjectifs comme *beau*, *mignon* et *gentil*. Dans la langue familière on a donc tendance à simplifier et à raccourcir le discours. Au lieu de chercher des comparants étendus et uniques comme c'est le cas en poésie ou en littérature d'une manière générale, on se contente d'utiliser une expression banale pour que le message à transmettre soit clair et pertinent au moment de l'énonciation.

Pour conclure, les comparaisons génériques, qui sont majoritaires parmi nos résultats, constituent le matériel linguistique pour nos prochaines analyses sémantiques

plus approfondies. Le tableau 4.4 représente donc le nombre des occurrences analysées en deuxième phase de notre recherche.

4.3 2^{ème} phase – « *être* + **ADJ** + *comme* + **SN** »

En deuxième phase, nous allons restreindre notre étude exclusivement à la comparaison générique à SN (cf. Tableau 4.4). Du point de vue fréquentiel, nous avons appris dans 4.2.5, que la comparaison figurative est assez abondamment répartie dans nos sous-corpus, et ce notamment dans les textes littéraires. Nous allons donc aussi approfondir nos connaissances sur ce type de comparaisons au niveau sémantique.

4.3.1 Associations comparatives

Nous allons classer maintenant des phrases comparatives à sens figuré par rapport à la nature des deux comparandes. Pour le faire, nous nous sommes inspirés de la distribution de Kokelberg (1993, p. 72) qui propose les associations suivantes :

- a. Un être humain à un être humain
- b. Un être humain à un animal/végétal
- c. Une « partie » d'être humain à une chose
- d. Une idée abstraite à une idée concrète/une idée concrète à une idée abstraite
- e. Un être humain à une réalité (chose) naturelle/artificielle
- f. Une chose à un être humain

Nous avons essayé de suivre le schéma indiqué. Or, au fur et à mesure qu'avancait notre analyse, nous nous sommes rendus compte que ce classement ne suffirait pas, car au final nous avons jusque trente combinaisons différentes pour comparer deux réalités. Peut-être le plus difficile était de distinguer les réalités **concrètes** des réalités **abstraites**. Les plus problématiques étaient des noms désignant des phénomènes naturels tels que l'air, le vent, le ciel, etc. Nous avons décidé de les classer finalement sous l'étiquette du « concret » qui regrouperait tout ce qu'on peut percevoir par n'importe lequel de nos sens⁴⁹.

Un autre problème posent les « êtres spéciaux » comme *ange*, *Dieu*, *vampire*, *mort* ou *cadavre*. Nous avons fini par les concevoir comme des humains à part *Dieu* avec

⁴⁹ **Exemple** : « *Il est rapide comme le vent foudroyant* [...] » (roman, Lanzmann, J., *La Horde d'or*, 1994)

majuscule, considéré comme une réalité abstraite. Par contre, les *dieux* quelconques ont été définis comme êtres humains si l'on se tourne plutôt vers l'idée des dieux antiques⁵⁰.

Concernant les végétaux, nous y avons classé aussi des légumes et des fruits bien qu'on puisse les considérer déjà comme des « produits » et donc objets⁵¹.

Quant à une partie d'être humain, ce terme englobe pour nous non seulement les parties du corps « physiques », mais aussi les actes physiologiques comme par exemple *regard* et *souffle*⁵², ou le mot *corps* en tant que tel⁵³.

De même, le terme des « noms collectifs » nous a permis d'y regrouper tous les noms que ce soit à sens propre du terme, comme par exemple *une armée*, ou des noms désignant un lieu fermé ou ouvert où la présence des humains est sous-entendue. De tels noms sont par exemple *ville*, *campagne*, *rue*, *bar*, etc.⁵⁴.

Il y figure naturellement des noms plus complexes encore, à propos desquels il faut plus réfléchir sur leur nature sémantique. En effet, nous faisons allusion aux noms polysémiques comme *chemin*, qui peut être concret ou abstrait selon le contexte.

Le tableau 4.5 représente les occurrences classées d'après les associations a – f., les autres associations minoritaires que nous avons observées n'y sont pas incluses. En nous penchant sur le tableau, nous avons abouti aux constatations suivantes.

En premier lieu, l'association incontestablement la plus fréquente dans notre matériel est la comparaison des êtres humains tant de leur aspects physiques (36) que psychiques (37) à des réalités naturelles ou artificielles⁵⁵ (320 occ.)⁵⁶.

⁵⁰ **Exemple :** « *Elles sont pures comme des anges et orgueilleuses comme des démons* » (théâtre, Montherlant, H. De, *Port-Royal*, 1954)

Par cet exemple nous voulons expliquer notre démarche. Nous avons classé les *anges* comme humains, car il nous semble qu'ils existent sous forme de tant représentations visuelles qu'ils sont presque humanisés. Par contre, nous plaçons *démon* contre *Dieu* avec majuscule, considérés comme abstraits.

⁵¹ **Exemple :** « [...] *je suis blanche comme une endive* » (roman, Malet, L., *Le Soleil n'est pas pour nous*, 1949)

⁵² **Exemple :** « [...] *leur sourire est las comme un feu qui s'éteint*. » (poésie, Samain, A., *Le chariot d'or*, 1900)

⁵³ **Exemple :** « [...] *une femme dont le corps jeune est frais comme l'eau des fontaines* » (poésie, Guerin, Ch., *Le Cœur solitaire*, 1904)

Ce procédé métaphorique de nommer un ensemble par sa partie s'appelle la **métonymie**. On pourrait objecter que le mot *corps* fait déjà référence à un homme, mais pour que celui-ci soit « complet », il lui manque encore le côté psychique ou l'âme.

⁵⁴ **Exemples :** « *Le restaurant en effet était plein comme un temple aux époques de foi miraculeuse.* » (roman, Hamp, P., *Marée fraîche, Vin de champagne*, 1909) ; « *Le village est inerte comme un grand cadavre étendu.* » (roman, Genevoix, M., *Ceux de 14*, 1950)

⁵⁵ Nous n'avons pas étudié la distribution proportionnelle exacte des réalités naturelles et artificielles, mais en gros les comparants artificiels étaient plus fréquents.

⁵⁶ Nous introduisons les pourcentages des occurrences de tous les corpus ainsi que les fréquences absolues des occurrences des sous-corpus particuliers.

(36) « Elle était **maigre** comme **une bicyclette** [...] » (théâtre, Pagnol, M., *Fanny*, 1932)

(37) « Je suis **triste** comme **une lampe à pétrole allumée en plein soleil** [...] » (roman, Cohen, A., *Le Livre de ma mère*, 1954)

Quant aux propriétés psychiques ou physiques, aucune tendance particulière pour le choix d'un comparant artificiel ou naturel ne se manifeste pas. D'ailleurs, c'est le sous-corpus des sous-titres (47 occ.) et du théâtre (48 occ.) dans lesquels l'association de l'animé à l'inanimé est la plus répandue. Cela pourrait s'expliquer par le fait que les deux domaines mentionnés sont conçus pour être mis en scène et que dans la langue orale, il est plus naturel d'utiliser cette sorte de comparaisons dite « matérielle »⁵⁷.

A l'inverse, la comparaison des réalités naturelles ou artificielles à des humains se rencontre rarement dans nos occurrences :

(38) « **Le ciel** était doux comme **une femme** [...] » (roman, Sartre, J-P., *La Mort dans l'âme*, 1949)

sous-corpus	humain & humain	humain & anim./vég.	partie hum. & chose	concret & abstrait	humain & chose	chose & humain
core – fiction	18	16	6	5	28	0
sous-titres	6	34	3	3	47	0
frWaC	7	5	1	7	12	1
L'Est républicain	7	11	0	6	23	0
PressEurop, Syndicate	0	0	0	3	0	0
Europarl	0	1	0	5	0	0
roman	66	128	41	39	129	3
poésie	13	19	10	14	24	1
théâtre	23	28	3	2	48	0
traités, essais	5	2	4	2	9	0
f absolue au TOTAL	145	244	68	86	320	5
N° de %	12 %	20 %	6 %	7 %	26 %	0,4 %

Tableau 4.5 : Fréquences absolues des associations comparatives dans chaque sous-corpus sélectionné de *KonText* et de *FRANTEXT* et fréquences absolues et pourcentages des associations comparatives dans tous les sous-corpus sélectionnés de *KonText* (excepté l'*Acquis communautaire* et *core – nonfiction*) et de *FRANTEXT*

En second lieu, le nombre assez important des comparaisons proviennent de l'association *être humain et animal/végétal* (20 %), utilisée le plus dans le roman (128 occ.)⁵⁸, les sous-titres (34 occ.) et le théâtre (28 occ.). Malgré le caractère hétérogène

⁵⁷ Nous généralisons ici en excluant des pièces de théâtre destinées uniquement à la lecture.

⁵⁸ Nous n'avons pas étudié la typologie des romans particuliers.

de ces associations, nous pouvons y rencontrer beaucoup de comparants **clichés**. En effet, associer un homme à un cochon ou une fille à une rose nous paraît peu exceptionnel. Un autre exemple est dans le sous-titre suivant (39), où comparer quelqu'un d'entêté à une mule n'a rien d'inhabituel.

(39) « *Tu es **têtu** comme une **mule**.* » (sous-titres, *The Legend of Zorro*, 2005)

D'ailleurs, déjà dans 4.2.1 et 4.2.2 nous avons émis l'hypothèse que les comparaisons clichées apparaissent souvent dans les sous-titres et les textes dramatiques.

En troisième lieu, l'association de l'être humain à un être humain (12 %) est sur la troisième place en matière fréquentielle. Les comparants humains sont souvent désignés dans ce cas par des noms propres⁵⁹.

En principe, plus les entités sont éloignées et rarement associées, plus la comparaison est originale. Faire référence à deux humains ou comparer un humain à un animal ne s'avère pas être très sophistiqué. A moins qu'il ne s'agisse d'un comparant vraiment inouï :

(40) « *Je suis sourd comme **un Soupault** et muet comme **un cachalot*** » (poésie, Soupault, P., *Georgia, Épitaphes, Chansons et Poèmes (1917 – 1983)*, 1983)

Il faut souligner que les textes contenus dans *FRANTEXT*, dont surtout les textes poétiques, sont des sources des comparaisons métaphoriques beaucoup plus variées et complexes en comparaison avec les autres corpus (cf. le SN étendu dans 4.2.4). En effet, les textes de *FRANTEXT* nous procurent des associations complètement nouvelles, par exemple en rapprochant une partie d'être humain à une chose abstraite (41) ou en utilisant le procédé de **défigement** (42) (cf. 3.2.5) :

(41) « [...] *vos yeux ignorants du rêve sont naïfs comme le bleu passé des faïences anciennes.* » (poésie, Guérin, CH., *Le cœur solitaire*, 1904)

(42) « *Voici les hommes*

Ils sont pâles comme des vivants » (poésie, Soupault, P., *Odes 1930-1980*, 1980)⁶⁰

⁵⁹ **Exemple** : « [...] *je serai riche comme Crésus* » (théâtre, Chepfer, G., *Saynetes, Paysanneries 2*, 1945)

Nous n'avons pas étudié la distribution proportionnelle des noms communs et propres.

⁶⁰ Normalement on dit « être pâle comme un mort » qui veut dire « être très pâle » (TLFi) [En ligne] [Consulté le 07.05.2016].

Bien sûr, il faut prendre en compte le caractère spécifique de la poésie surréaliste.

D'un côté, la caractéristique commune à l'instar de laquelle on met en rapport de comparaison deux réalités est dans (41) très opaque. D'un autre côté, l'actualisation d'une locution figée (42) est, effectivement, un des moyens utilisés pour créer des comparaisons originales.

En conclusion, dans l'ensemble il faut dire qu'il était peu aisé de détecter une tendance globale pour le choix des comparandes à cause de la grande hétérogénéité des occurrences⁶¹. Cela dit, il fait reconnaître que les textes littéraires reflètent un inventaire d'imagination illimité.

Notre intention dans les prochains sous-chapitres est de nous consacrer justement à des associations peu originales qui figurent déjà dans les dictionnaires comme figées ou qui soient fréquentes parmi nos résultats.

4.3.2 Adjectifs « candidats » au figement

Parmi les comparaisons figuratives étudiées, nous en avons repéré quelques-unes qui sont répétées plusieurs fois. En effet, la fréquence est un des moyens pour détecter une locution figée. *Le dictionnaire des expressions et locutions* (Rey & Chantreau, 1991, p. 231) affirme que *comme* joue un rôle fondamental dans la phraséologie française⁶².

Pour commencer, nous avons bénéficié de l'outil de *KonText* qui affiche des listes de fréquences. Ainsi, nous avons créé trois listes fréquentielles des adjectifs les plus fréquents apparaissant dans la construction « être+ ADJ + *comme* » à partir des corpus *InterCorp* du français dans son ensemble⁶³, *L'Est républicain* et *frWaC*⁶⁴.

Étant donné le nombre de résultats peu élevé, nous avons décidé de déterminer la limite de fréquence minimale des adjectifs au moins à cinq. De cette façon, nous avons sélectionné **quatorze adjectifs candidats** à condition que ces derniers se trouvent sur les dix premiers rangs au moins dans une de nos listes fréquentielles. Les adjectifs choisis sont *beau, heureux, doux, simple, blanc, noir, frais, vieux, solide, plein, sage, froid, long* et *bleu*

Un autre exemple en pourrait être « [...] *il était haut comme une pomme* [...] » (roman, Sarraute, N., *Le Planétarium*, 1959), car il existe une locution figée « *haut comme trois pommes* » qui signifie paradoxalement « être tout petit ». (TLFi) [En ligne] [Consulté le 07.05.2016].

⁶¹ Les comparaisons à sens péjoratif ou mélioratif sont présentes les deux sans que nous ayons repéré des tendances particulières.

⁶² Par le mot *phraséologie* nous comprenons ici l'ensemble des tournures typiques d'une langue par leur fréquence ou par leur caractère idiomatique (TLFi) [En ligne] [Consulté le 30.04.2016]

⁶³ Nous avons visualisé la liste fréquentielle des adjectifs dans l'ensemble du corpus pour avoir un nombre de résultats significatif, mais en excluant le sous-corpus *Acquis communautaire*, éliminé de notre recherche déjà dans 4.2.3 à défaut de structures comparatives.

⁶⁴ Dans ce cas *FRANTEXT* ne va pas nous servir, car il permet de faire seulement des listes fréquentielles de mots isolés.

se rapportant tant aux aspects physiques que psychiques. Nous pouvons les voir rangés selon leurs fréquences dans le tableau ci-dessous :

	InterCorp		L'Est républicain		frWaC	
	lemme	f abs.	lemme	f abs.	lemme	f abs.
1.	être beau comme	63	être simple comme	31	être beau comme	584
2.	être juste comme	33	être heureux comme	26	être fou comme	579
3.	être blanc comme	32	être fou comme	20	être bleu comme	326
4.	être doux comme	31	être présent comme	19	être simple comme	292
5.	être fou comme	29	être beau comme	16	être vieux comme	224
6.	être grand comme	25	être plein comme	13	être présent comme	170
7.	être heureux comme	21	être vieux comme	11	être sympa comme	170
8.	être drôle comme	20	être frais comme	11	être marrant comme	151
9.	être simple comme	16	être sage comme	9	être heureux comme	147
10.	être froid comme	15	être solide comme	8	être humain comme	145
11.	être noir comme	15	être ouvert comme	8	être perçu comme	139
12.			être drôle comme	8		
13.			être incroyable comme	7		
14.			être longue comme	7		

Tableau 4.6 : Dix premiers rangs dans les listes fréquentielles des corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC* avec les adjectifs candidats au figement marqués en couleurs différentes et les fréquences absolues de ces derniers⁶⁵

Comme nous pouvons le voir dans le tableau 4.6, aux premiers rangs apparaissent aussi les adjectifs « bruisants », analysés dans le sous-chapitre 4.2.3 ainsi que ceux liés à un autre contexte droit que nominal (cf. 4.2.4)⁶⁶. En ce qui concerne le corpus *FRANTEXT*, dans chacun de nos sous-corpus créés nous avons filtré les adjectifs candidats sélectionnés dans les corpus précédents. Nous avons pris en compte également le genre féminin et le pluriel. A part *frWaC* qui est, quant à lui, le corpus le plus vaste en positions, mais avec une variabilité lexicale peu importante (cf. Graphique 4.2 dans 4.2.1), les fréquences

⁶⁵ Les adjectifs de la même fréquence partagent pour nous un rang. C'est pourquoi nous avons choisi des adjectifs allant jusqu'au rang n°14 dans les corpus.

⁶⁶ Concrètement les adjectifs *juste* dans un emploi adverbial ; *fou*, *drôle*, *incroyable* et *marrant* dans des phrases exclamatives ; *perçu* comme participe nécessitant une préposition ; *humain* lié à au substantif *être* ; *sympa* utilisé notamment pour exprimer l'appartenance à une classe et puis suivi du pronom *tout* ; *grand* suivi dans la majorité des cas du pronom *ça* ; *présent* suivi principalement de circonstants, par ex. « [...] *ses deux filles et son gendre, étaient présents comme chaque année à cette cérémonie* [...] » (*L'Est républicain*, 25.08.2002) ; et *ouvert* employé seulement dans la séquence à SV suivante : « *ouverte comme suit* : [...] » (*L'Est républicain*).

absolues ainsi que les pourcentages relatifs de tous les adjectifs candidats figurant dans la construction avec *comme* sont assez réduites :

adjectif	core - f.	%	sous- titres	%	PressE. & Synd.	%	Euro- parl	%	L'Est rép.	%	frWaC	%
beau	20	10%	42	6%	1	5%	0	0%	16	3%	584	5%
heureux	8	4%	12	2%	0	0%	1	4%	26	5%	147	1%
doux	3	1%	27	4%	0	0%	0	0%	2	0%	137	1%
simple	2	1%	13	2%	1	5%	0	0%	31	6%	292	2%
blanc	5	2%	27	4%	0	0%	0	0%	5	1%	136	1%
noir	5	2%	9	1%	0	0%	0	0%	1	0%	70	1%
frais	0	0%	4	1%	0	0%	0	0%	11	2%	27	0%
vieux	1	0%	2	0%	1	5%	1	4%	11	2%	224	2%
solide	0	0%	7	1%	2	10%	1	4%	8	1%	25	0%
plein	1	0%	1	0%	0	0%	0	0%	13	2%	50	0%
sage	0	0%	2	0%	0	0%	0	0%	9	2%	50	0%
froid	3	1%	12	2%	0	0%	0	0%	1	0%	29	0%
long	0	0%	2	0%	0	0%	0	0%	7	1%	36	0%
bleu	0	0%	3	0%	0	0%	0	0%	3	1%	326	3%

Tableau 4.7 : Fréquences absolues et pourcentages des adjectifs candidats au figement dans la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *KonText* (excepté l'Acquis communautaire)

adjectif	roman	%	poésie	%	théâtre	%	traités, essais	%
beau	66	5%	11	6%	10	4%	8	3%
heureux	36	3%	4	2%	5	2%	3	1%
doux	25	2%	2	1%	6	3%	1	0%
simple	23	2%	2	1%	3	1%	1	0%
blanc	24	2%	2	1%	3	1%	1	0%
noir	23	2%	5	3%	0	0%	1	0%
frais	10	1%	2	1%	2	1%	3	1%
vieux	9	1%	0	0%	2	1%	3	1%
solide	11	1%	0	0%	1	0%	0	0%
plein	3	0%	1	1%	2	1%	1	0%
sage	3	0%	1	1%	4	2%	0	0%
froid	10	1%	0	0%	1	0%	0	0%
long	5	0%	3	2%	2	1%	1	0%
bleu	13	1%	3	2%	1	0%	0	0%

Tableau 4.8 : Fréquences absolues et pourcentages des adjectifs candidats au figement dans la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *FRANTEXT*

Nous pouvons en déduire que plus le taux de variabilité lexicale d'un sous-corpus est grand, moins de résultats nous avons pour les adjectifs candidats au figement. Ceci peut découler bien sûr de la construction dans laquelle nous les avons cherchés, car elle-même est très rare dans tous les corpus consultés (cf. Graphique 4.1 dans 4.2.1).

De plus, les listes de fréquence dans tous les sous-corpus de *KonText* nous font découvrir l'existence de nombreux « **hapax legomena** » dans la construction « *être* + ADJ + *comme* », c'est-à-dire les lemmes qui n'ont qu'un seul token.

Afin d'examiner si les premiers rangs des adjectifs candidats au figement dans la construction ne sont que corrélés à la haute fréquence des adjectifs sans contexte spécifié, nous avons cherché individuellement chaque adjectif dans l'ensemble des corpus *InterCorp*, *L'Est républicain* et *frWaC*. Dans le tableau qui suit, nous avons inséré les fréquences relatives des adjectifs cherchés comme lemmes :

InterCorp	ipm	L'Est républicain	ipm	frWaC	ipm
beau	320,79	beau	401,91	beau	278,79
vieux	268,65	plein	198,61	simple	210,5
plein	196,76	long	164,65	plein	187,37
long	196,15	blanc	115,22	long	169,64
heureux	150,42	heureux	103,56	blanc	138,64
noir	139,63	simple	101,5	noir	128,41
simple	134,46	vieux	92,67	frais	126,43
blanc	113,47	noir	88,99	vieux	88,02
froid	81,67	bleu	85,24	bleu	66,60
bleu	62,42	frais	47,76	froid	60,95
doux	54,19	froid	44,82	doux	56,32
frais	40,02	solide	30,95	heureux	54,12
solide	38,26	doux	28,99	solide	35,98
sage	25,69	sage	11,13	sage	16,52

Tableau 4.9 : Fréquences relatives (ipm) des adjectifs candidats au figement hors la construction « *être* + ADJ + *comme* » dans les corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC*

Comparons maintenant l'ordre des adjectifs dans les tableaux 4.6 et 4.9. Il est évident que l'adjectif *beau* prédomine quant au nombre d'occurrences, qu'il figure hors ou dans la construction comparative. Par conséquent, nous pouvons dire que pour certains adjectifs comme *beau* ou *simple*, leur haute fréquence dans la construction étudiée est le reflet de la haute fréquence de ces mots dans l'ensemble du corpus.

Par contre, les adjectifs peu fréquents dans la construction, tels que *plein*, *vieux*, *noir* ou *long* figurent seuls sur les premiers rangs. Finalement, les fréquences relatives de certains adjectifs qui apparaissent aux dix premiers rangs dans la construction, par exemple *solide*, *doux* ou *sage*, sont peu marquantes sans contexte déterminé. Cela pourrait être un bon signe pour notre recherche des locutions figées issues de la construction « être + ADJ + *comme* » dans le dernier sous-chapitre à venir.

4.3.3 Locutions figées adjectivales avec *comme*

Dans 3.2.5 nous avons présenté les caractéristiques principales des locutions figées. Celles-ci peuvent être formées par toutes les classes des mots dont les adjectifs sont la source de locutions très nombreuses et les plus variées (Čermák & Šulc, 2006, p. 142).

Dans un premier temps, afin de découvrir des locutions figées dans les corpus utilisés, nous allons chercher des locutions formées par les adjectifs candidats au figement (cf. 4.3.2) dans les dictionnaires. Nous allons nous servir principalement du *Dictionnaire d'expressions idiomatiques*⁶⁷, du *TLFi*⁶⁸ et du *Dictionnaire des expressions et locutions* (Rey & Chantreau, 1991).

Dans un second temps, nous allons utiliser la fonction « collocations »⁶⁹ de *KonText* dans le but de chercher d'autres locutions potentiellement figées lesquelles ne figurent pas dans les dictionnaires comme lexicalisées. Si nous en trouvons quelques-unes, des tests linguistiques vont être appliqués pour pouvoir déterminer leur degré de figement et peut-être découvrir de nouvelles locutions figées.

Dans le tableau 4.10 ci-dessous sont présentées toutes les locutions basées sur les adjectifs candidats qui figurent déjà comme lexicalisées dans les dictionnaires consultés. Beaucoup d'entre elles ont été trouvées aussi dans les corpus (cf. Tableau 4.10). Comme nous pouvons le remarquer, le plus de locutions figées sont formées par les adjectifs *beau*, *blanc* et *frais*. D'ailleurs, les locutions avec *beau* figurent toutes dans les corpus. Bien évidemment, toutes ces locutions ne sont pas mentionnées dans tous les dictionnaires. C'est la raison pour laquelle nous avons été obligés de consulter plusieurs dictionnaires. En effet, il se peut tout simplement que le degré de figement ne soit pas un paramètre fixe et que certaines locutions soient plus communes que d'autres.

⁶⁷ [En ligne] [Consulté le 24.04.2016].

Disponible sur : <http://www.cnrtl.fr/dictionnaires/expressions_idiomatiques/>

⁶⁸ [En ligne] [Consulté le 24.04.2016]. Disponible sur : <<http://atilf.atilf.fr/tlf.htm>>

⁶⁹ Le terme « collocation » fait référence même à des locutions qui ne sont pas des idiotismes (cf. 3.2.3).

beau	un astre, un dieu, le jour, un ange, la lumière, un camion, un cœur
heureux	un roi, un pape, un poisson (dans l'eau)
doux	le miel, un agneau
simple	bonjour
blanc	un cachet d'aspirine, un cygne, neige , un lavabo, un linge , un drap, un cadavre, un mort
noir	du cirage , le diable, du charbon, de l'encre , une taupe
frais	le jour, une fleur, une pomme d'api, l'œil, une rose, un gardon , une pâquerette
vieux	les chemins, les rues, Hérode, Mathusalem, le monde
solide	un chêne, le Pont-Neuf , un bœuf, un roc
plein	un boudin, une barrique, un œuf , une outre, une huître
sage	une image , une fille
froid	un/du/le marbre, une carafe d'orgeat
long	un jour sans pain
bleu	X

Tableau 4.10 : Mots entrant en locutions figées avec *comme* et les adjectifs candidats au figement dans la construction « être + ADJ + *comme* », trouvés dans les dictionnaires utilisés; mots trouvés aussi dans les corpus sont marqués en gras

Les adjectifs candidats sont associés à des locutions figées très variées. La situation est similaire pour des locutions figurant dans les corpus y compris *FRANTEXT*, sauf que la diversité y est encore plus marquante. Nous proposons donc de voir quelques mots se trouvant en collocation avec *comme* et les adjectifs candidats que nous avons rencontré dans les corpus pour pouvoir en déduire des tendances associatives globales.

Nous avons tenté de donner un peu d'ordre aux associations trouvées en classant certains adjectifs selon les champs lexicaux desquels ceux-là se rapprochent. Dans le tableau 4.11 nous pouvons voir par exemple, que ce qui est typique pour l'expression de la beauté, c'est la comparaison au soleil et aux dérivés de ce dernier, pour le bonheur la comparaison à un enfant et à ses synonymes, pour *vieux* que celui-ci fait référence surtout à l'origine biblique, pour l'adjectif *froid* sa connotation à l'hiver et aux dérivés de celui-ci, et en fin de compte, l'adjectif *long* est perçu surtout comme abstrait en se rapportant à l'idée du temps.

	beau
FRANTEXT	humain : un prince, une femme œuvre d'art : une belle image, une œuvre d'art, un tableau soleil : une constellation, les aurores d'été, l'aube, un coucher de soleil
KonText	humain : un prince œuvre d'art : un tableau soleil : le soleil, un coucher de soleil, un rayon de soleil autres : un sou neuf, des fleurs, l'antique
	heureux
FRANTEXT	enfant : un/l'enfant/un petit enfant candide, un petit d'animal humain : un héros de roman, un mort autres : des lentilles dans l'eau, un pou
KonText	enfant : un gamin, un même, un gosse, un petit garçon humain : un niais, un fou, un prince animal : un pinson, un coq en pâte, un pou, un cochon
	doux
FRANTEXT	animal : une agnelle, un mouton, un toutou phénomène naturel : la chute des flocons de neige, la pluie aliment : le sucre, l'eau pure, le lait autres : un morceau de soie, une femme, un baume, la soif apaisée, la musique de flûtes
KonText	animal : un mouton, un chaton, un ourson aliment : de la crème fraîche, la grenade phénomène naturel : la brise autres : un bain chaud, du duvet, du coton
	simple
FRANTEXT	une bête
KonText	? : un coup de fil, un jeu d'enfant, un clic, la vie, un moteur de recherche, un coup de bambou, bonsoir la clim, l'oeuf de Christophe Colomb
	blanc
FRANTEXT	partie d'animal : une barbe de chèvre, le ventre de la brebis, la laine objet : de la cire, une feuille de papier, des berceaux, la neige/neige, un paysage de neige aliment : une endive, un navet, un sucre autres : la lumière, une mariée
KonText	objet : du papier, la craie le corps, l'agneau, du chrome, un navet
	noir
FRANTEXT	mort : une nuit, l'enfer humain : des mineurs, des négresses, ses pieds objet : jais, du café, un four, les pots d'encre, une bourrique, des pruneaux, des jus de pruneau, les mines, le plomb
KonText	mort : la mort, la nuit, l'enfer humain : un charbonnier autres : jais, l'abysse, l'ébène, un corbeau, un four, un linceul
	frais
FRANTEXT	aliment : un vin, une pomme, une praline jardin : une rosée, une petite rose, l'eau des fontaines, une ombrelle autres : la nuit

KonText	air : air de campagne, le mois d'avril, un matin d'hiver
	vieux
FRANTEXT	origine (biblique) : Dieu, ce vieux monde, le premier homme, l'antéchrist, l'humanité, la civilisation
KonText	origine (biblique) : la Bible, le diable, Adam, l'antiquité, le péché politique : la politique, la corruption, le PS
	solide
FRANTEXT	plante : les grands chênes, la terre
KonText	construction : la tour du roi, un fondeur, l'acier, un rocher
	plein
FRANTEXT	un tourteau
KonText	une ruche, un gros mélange de coton sale
	sage
FRANTEXT	X
KonText	un ange, un petit mouton
	froid
FRANTEXT	hiver : de la glace, la neige autres : la bouche d'un mourant, un poisson, un fourneau éteint, le ciel
KonText	hiver : la glace, un iceberg, du verglas, l'hiver matériel : du métal, de la pierre, les marbres, l'acier, le grés, une lame mort : le temps, la mort
	long
FRANTEXT	couloir, une carabine, assaut, deux petits poissons, une lamproie, un long gant de bal
KonText	temps : l'horizon, long comme une année (le jour), l'infini, un jour sans fin, la vie, un jour sans soleil sommeil concret : un doigt, le bras, un crayon
	bleu
FRANTEXT	concret ⁷⁰ : une orange, le voile de la Vierge, un drapeau, le ciel, de l'eau profonde abstrait : une fête, le jour, le soir, un songe
KonText	aliment : une orange, une tomate élément naturel : la comète, la planète, le ciel/un ciel d'été, l'océan, les îles Caraïbes autres : œil plein de larmes, l'enfer, un doux rêve d'enfant

Tableau 4.11 : Mots entrant en locution (autre que figée) avec *comme* et les adjectifs candidats au figement dans la construction « être + ADJ + comme » dans les corpus sélectionnés de *KonText* (*InterCorp* excepté l'*Acquis communautaire*) et de *FRANTEXT*

Maintenant, nous allons essayer de trouver des locutions figées en nous penchant sur une méthode statistique. Concrètement, nous allons utiliser les **mesures associatives** dont *KonText* nous fait bénéficier⁷¹. Les mesures les plus utilisées sont **MI-score** (ou *MI*), **T-score** et **logDice** (Čermák & Šulc, 2006, p. 13)⁷². Sur **KWIC**⁷³, nous bornons l'échelle du

⁷⁰ La limite abstrait/concret peut être différente dans ce tableau par rapport à ce qu'elle est dans 4.3.1.

⁷¹ Dans *FRANTEXT* il est possible d'étudier le voisinage d'un mot aussi, mais il faut noter que ce corpus ne travaille pas avec des mesures associatives.

⁷² Chacune de ces mesures fonctionne d'une manière un peu différente si bien qu'on utilise l'une ou l'autre conformément à ce dont on a besoin de chercher (Čermák & Šulc, 2006, p. 13). Nous utilisons les trois et

contexte gauche à 0 mot et le contexte droit à 2 mots⁷⁴. La fréquence minimale des lemmes de la séquence dans l'ensemble d'un corpus est déterminée à 3 et la fréquence minimale de la séquence donnée à 3 aussi. Nous cherchons *KWIC* dans les trois corpus de *KonText*, dans *InterCorp* de nouveau dans son intégralité en excluant l'*Acquis communautaire*.

Après avoir effectué la recherche, nous constatons que *T-score* n'est pas applicable à notre étude. Les lemmes sur les premiers rangs sont surtout des mots grammaticaux comme les déterminants ou les pronoms personnels, très fréquents en général. En effet, cette mesure est plus utile si l'on a des lemmes beaucoup plus fréquents (Čermák & Šulc, 2006, p. 13).

D'après le tableau 4.12 ci-dessous nous pouvons constater que, concernant *MI*⁷⁵ et *logDice*⁷⁶, la disposition de leurs rangs est assez semblable. La seule différence majeure est observable dans le corpus *frWaC* dans lequel la mesure *MI* n'affiche rien qui pourrait nous servir sur les premiers rangs. Les locutions figées commencent néanmoins à apparaître petit à petit à partir du rang n°13. Dès lors, la taille importante du corpus joue un grand rôle et nous avons un indice que *MI* convient mieux à la recherche des collocations peu fréquentes (Ibid.).

rang	InterCorp logDice	f	T- score	MI	logDice	rang	InterCorp MI	f	T- score	MI	logDice
1.	roc	8	2.828	12.542	7.866	1.	barrique	4	2.000	13.750	7.007
2.	agneau	10	3.160	10.795	7.722	2.	pinson	4	2.000	13.750	7.007
3.	neige	16	3.996	9.871	7.603	3.	gardon	3	1.732	12.801	6.574
4.	linge	10	3.160	10.485	7.602	4.	berceuse	3	1.732	12.669	6.568
5.	carpe	5	2.235	11.529	7.143	5.	roc	8	2.828	12.542	7.866
6.	mule	5	2.235	10.976	7.048	6.	carpe	5	2.235	11.529	7.143
7.	barrique	4	2.000	13.750	7.007	7.	mule	5	2.235	10.976	7.048
8.	pinson	4	2.000	13.750	7.007	8.	braise	3	1.731	10.853	6.415
9.	chaton	4	1.999	10.660	6.727	9.	agneau	10	3.160	10.795	7.722
10.	gardon	3	1.732	12.801	6.574	10.	chaton	4	1.999	10.660	6.727

d'après l'ordre des rangs qu'elles afficheront, nous pourrions décider quelle suit le mieux à notre étude, éventuellement nous allons les combiner.

⁷³ Key word in context = séquence cherchée dans le contexte

⁷⁴ Parfois nous pouvons découvrir des collocations jusqu'à la distance de cinq mots.

⁷⁵ MI= mutual information

Cette mesure est basée sur la probabilité de découverte des éléments *x* et *y* ensemble ainsi que sur la probabilité de découverte de l'élément *x* et de l'élément *y* indépendamment l'un de l'autre.

⁷⁶ A la différence de *MI*, *logDice* ne prend pas en compte la taille du corpus.

rang	L'Est rép. logDice	f	T-score	MI	logDice	rang	L'Est rép. MI	f	T-score	MI	logDice
1.	gardon	9	2.999	12.277	8.157	1.	gardon	9	2.999	12.277	8.157
2.	sou	5	2.235	10.772	6.964	2.	sou	5	2.235	10.772	6.964
3.	œuf	10	3.158	9.535	6.304	3.	gosse	3	1.730	10.099	6.264
4.	gosse	3	1.730	10.099	6.264	4.	œuf	10	3.158	9.535	6.304
5.	orange	3	1.728	8.773	5.368	5.	orange	3	1.728	8.773	5.368
6.	poisson	8	2.820	8.364	5.220	6.	poisson	8	2.820	8.364	5.220
7.	pomme	4	1.990	7.623	4.458	7.	pomme	4	1.990	7.623	4.458
8.	témoigner	4	1.989	7.531	4.374	8.	Il	3	1.723	7.541	4.341
9.	Il	3	1.723	7.541	4.341	9.	témoigner	4	1.989	7.531	4.374
10.	image	8	2.805	6.937	3.866	10.	image	8	2.805	6.937	3.866

rang	frWaC logDice	f	T-score	MI	logDice	rang	frWaC MI	f	T-score	MI	logDice
1.	orange	279	16.689	10.162	7.700	1.	bbbbrrrrrrrrrr	3	1.732	16.981	2.978
2.	agneau	34	5.818	8.797	5.639	2.	repos.de	5	2.236	16.718	3.714
3.	rétractilité	17	4.123	16.609	5.478	3.	rétractilité	17	4.123	16.609	5.478
4.	braise	20	4.468	10.139	5.472	4.	porçin	3	1.732	16.566	2.978
5.	chou	32	5.636	8.086	5.242	5.	drv	4	2.000	14.337	3.390
6.	hab	20	4.459	8.435	5.038	6.	botin	4	2.000	13.981	3.390
7.	neige	61	7.762	7.343	5.032	7.	bronchitique	3	1.732	13.866	2.975
8.	roc	15	3.867	9.447	5.010	8.	tulipa	3	1.732	13.708	2.975
9.	carpe	17	4.114	8.785	4.996	9.	Obligé	4	2.000	13.148	3.387
10.	garde-fou	13	3.603	10.265	4.944	10.	calzone	3	1.732	12.981	2.973

Tableau 4.12 : Lemmes entrant en collocation avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans les corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC* ; avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*

De plus, il est à noter que dans tous les corpus, les dix premiers rangs comprennent, à part quelques exceptions, des locutions figées y compris des locutions figées liées aux adjectifs candidats trouvées dans les dictionnaires⁷⁷.

⁷⁷ Concrètement « *solide comme un roc* », « *doux comme un agneau* », « *blanc comme neige* », « *plein comme une barrique/un œuf* », « *frais comme un gardon* », « *sage comme une image* » et « *heureux comme un poisson dans l'eau* »

Cependant, les deux mesures, *MI* comme *logDice*, interprètent comme locutions figées également des locutions issues de l’idiolecte d’un auteur. Ainsi, malgré le nombre d’occurrences peu important, la séquence « *beau comme **une berceuse*** » apparue trois fois chez Kundera figure sur le rang n°4 dans les belles-lettres (*MI*) et la séquence « *Il est beau beau comme la rétractilité des serres des oiseaux rapaces [...]* », toujours la même, figure sur le rang n°3 dans *frWaC* (17 occ.).

De la même manière, la séquence que les mesures ont identifiée comme collocation sur le 1^{er} rang dans *frWaC* (*logDice*) et sur le rang n°5 dans *L’Est républicain* est le fameux premier vers du poème de Paul Eluard, à savoir « *La terre est bleue comme **une orange*** ». En effet, tous les résultats ne font que citer ce vers. Dans *FRANTEXT* nous en avons détecté une occurrence dans le roman sans qu’elle fasse référence à son inventeur. Cela pourrait signifier que certains auteurs reprennent la locution en la considérant déjà comme un trope (cf. 3.2.4).

Dans l’intention de découvrir de nouvelles locutions potentiellement figées, il est préférable de chercher un peu plus loin dans les listes. Dans Čermák (2007, p. 206) il est mentionné que la valeur- limite de *MI-score* pour les locutions figées tourne autour de la valeur 7. Si la valeur est inférieure, il s’agit le plus souvent des combinaisons des adjectifs fréquents et des noms fréquents (Ibid.). Néanmoins, *MI* dans nos listes est de valeur inférieure pour beaucoup de locutions figées (cf. Annexe C). La cause en peuvent être les faibles fréquences des locutions dans les corpus⁷⁸. C’est la raison pour laquelle il faut filtrer les listes de collocations plutôt manuellement⁷⁹.

locution	rang	InterCorp MI	f	rang	L’Est rép. MI	f	rang	frWaC MI	f	f FRANTEXT
froid comme la glace	22.	7.953	6	X	X	0/0	333.	3.771	3/4	1 (de la)
simple comme un coup de fil	X	X	1	15.	6.417	12/14	230.	4.646	43/55	0
doux comme un mouton	X	X	2/31	X	X	0/2	130.	5.992	2/8	3

Tableau 4.13 : Locutions avec *comme* et les adjectifs candidats au figement, à savoir *froid*, *simple* et *doux*, issues de la construction « être + ADJ + *comme* » dans les corpus sélectionnés de *KonText* (*InterCorp* excepté l’*Acquis communautaire*) et de *FRANTEXT*, choisies comme potentiellement figées ; avec l’affichage du n° de rang dans la liste de collocations, de la valeur de la mesure *MI* et de la fréquence absolue dans le corpus

⁷⁸ Dans l’Annexe C nous affichons les rangs jusqu’au rang n° 39 pour *InterCorp*, car c’est là où se trouve la limite approximative pour des locutions figées, pour *L’Est républicain* jusqu’au rang n°30 et *comme frWaC* est très dispersé et des locutions figées se trouvent même sur des rangs plus éloignés, nous introduisons au moins 50 rangs comme illustration.

⁷⁹ Bien sûr, le mieux est quand le filtrage est effectué par un locuteur natif.

Nous avons combiné les mesures associatives avec les listes fréquentielles et nous avons fini par choisir trois locutions figées potentielles figurant dans le tableau 4.13.

Premièrement, la locution qui nous semble potentiellement figée est « *doux comme un mouton* ». Bien que les mesures associatives ne lui attribuent pas le caractère figé, nous l'avons paradoxalement trouvé dans un dictionnaire, à savoir dans *TLFi*. En effet, si nous faisons la recherche sur « *doux comme* », la locution n'y figure pas. En revanche, lorsque nous avons cherché déjà la séquence « *comme un mouton* », la locution a été affichée, mais sans aucun commentaire s'il s'agit d'une locution figée ou non. Nous supposons donc qu'il est question soit d'une locution figée à la limite soit d'une locution défigée, formée à partir de la locution « *doux comme un agneau* » (cf. Tableau 4.10).

Deuxièmement, nous nous doutons du caractère figé de la locution « *c'est simple comme un coup de fil* » dans son intégralité. Celle-ci apparaît dans les trois corpus de *KonText*. Dans le corpus *InterCorp* la locution ne compte qu'un seul résultat dans les sous-titres⁸⁰. Vu le Tableau 4.9, nous pouvons dire que l'adjectif *simple* est relativement fréquent parmi les adjectifs candidats, mais cela ne doit pas être le cas parmi tous les adjectifs du corpus. Concernant le tour *coup de fil*, nous soupçonnons qu'il ne soit pas très répandu. Par contre, le mot *coup* pourrait être assez fréquent. Bien entendu, non toutes les collocations avec *coup* sont étendues par *fil* (cf. Tableau 4.13). Parmi des locutions trouvées, quelques-unes pourraient être des locutions « défigées » comme par exemple « *simple comme un coup de clic* » (frWaC) ou « *simple comme un coup de téléphone* » (frWaC).

En ce qui touche l'opacité sémantique, elle ne se manifeste pas vraiment dans ce cas. Si quelque chose est simple comme un coup de fil, cela veut dire que c'est aussi simple que téléphoner, une activité assez confortable, rapide et qu'on a à disposition quasiment partout.

Pour savoir si la structure peut subir des transformations, nous utilisons des **tests linguistiques** repris en partie de Gross (1996, p. 12)⁸¹ et de Moline (2006, p. 19) en bas⁸² :

- a. **Reprise anaphorique** : ?C'est simple comme l'est un coup de fil⁸³.

⁸⁰ Nous avons essayé de changer la limite de la fréquence minimale à 1, mais la locution n'a pas été évaluée comme collocation dans ce corpus quand même.

⁸¹ Gross (1996, p. 12) cite encore autres tests linguistiques, à savoir la passivation et la relativation. Cependant, ces tests ne s'accordent pas avec la construction « *être + ADJ + comme* ».

⁸² Nous avons consultés les tests auprès de deux locuteurs natifs et nous introduisons les évaluations de chacun d'eux.

⁸³ « * » marque une phrase agrammaticale, « ?? » une phrase à la limite d'acceptabilité et « ? » une phrase sémantiquement acceptable

- b. **Insertion** (marqueur d'intensité) : *C'est très simple comme un coup de fil⁸⁴.
- c. **Négation** : ??Ce n'est pas simple comme un coup de fil.
- d. **Détachement** (extraction) : *Comme un coup de fil, c'est simple.
- e. **Actualisation** : C'est simple comme *le/ ?/*ce/ ??/*Ø coup de fil.
- f. **Métaphorisation** : ??C'est un coup de fil.
- g. **Substitution** : ?/*Il est simple comme un coup de fil. / ?/*Nous sommes simples comme des coups de fil.

? C'est facile comme un coup de fil.

En essayant de transformer la structure, nous nous rendons compte que les changements semblent peu naturels. En effet, les deux locuteurs natifs qui, à part la reprise anaphorique dans *a.* et la substitution de l'adjectif *simple* par son synonyme *facile* dans *g.*, considèrent les transformations comme étranges. D'ailleurs, la métaphore ne donne pas le sens tel qu'elle devrait donner. La locution paraît donc assez figée.

De plus, la moitié des occurrences dans *L'Est républicain* et presque 20 % des résultats dans *frWaC* représentent dans les textes différents la même phrase, à savoir « **Le bonheur, c'est simple comme un coup de fil** ». C'est pourquoi nous avons essayé de chercher la locution dans Google et nous avons découvert que la phrase vient d'une publicité de *Télécom* des années 80⁸⁵. L'origine de la phrase est même expliquée dans une des occurrences :

(43) « *N'a-t-on pas désormais l'habitude de faire vivre au quotidien le slogan - déjà obsolète - de France Télécom : « Le bonheur, c'est simple comme un coup de fil » ?* » (*L'Est républicain*, 27.09.2002)

La présence de la locution citée dans les textes tirés du web est donc claire. Par ailleurs, c'est à partir de cette phrase que sont formées des phrases défigées parmi les autres occurrences, comme par exemple « *L'amour/La santé, c'est simple comme un coup de fil* », toujours sous forme des slogans, propres au domaine journalistique. Sinon, le domaine dans lequel nous avons noté le plus d'occurrences est le domaine téléphonique :

(44) « *Trouver celui qui vous convient? C'est simple comme un coup de fil! Téléphone sans fil avec technologie DECT La DECT (Digital Enhanced*

⁸⁴ En effet, la comparative en *comme* et le modifieur de l'adjectif occupent le même paradigme. Pour que la phrase devienne grammaticalement correcte, il faudrait y ajouter une virgule : « ?? *C'est très simple, comme un coup de fil.* ». L'acceptabilité est cependant une autre chose.

⁸⁵ Vidéo de la publicité est disponible sur : <<https://www.youtube.com/watch?v=ki4TLe2EGAo>>

Cordless Telecommunication [sic]) est une technologie française de transmission numérique par liaison radio. » (frWaC)

Troisièmement, quant à l'opacité sémantique de la locution « **froid comme la glace** », de la même façon que pour la précédente, du point de vue sémantique la locution est assez transparente. Ce qui peut varier, c'est son emploi à l'égard d'une propriété physique ou d'un trait de caractère :

(45) « [sic] *les gens a mon gout ne sont pas tres sympas attention je m'explique il y a pleins de personnes bien !! ce que je veux dire c'est qu'il y en a qui sont froid comme la glace et moi j'en ai trop marre de l'eure et loir et compagnie !* »⁸⁶ (frWaC)

(46) « *Car, même à travers le tissu froid et imbibé d'eau de sa chemise, il sentit que ces deux mains étaient encore plus froides – qu'elles étaient froides comme la glace* » (belles-lettres, Bulgakov, M., *Le maître et Marguerite*⁸⁷)

Parmi nos occurrences, 7 résultats sur 9 sont envisagés avec un sens figuré pour qualifier le comportement d'une personne (45).

Puis, essayons de faire des tests linguistiques :

- a. **Reprise anaphorique** : ?Il est froid comme l'est la glace.
- b. **Insertion** : *Il est très froid comme la glace.
?Il est très froid, comme la glace.
- c. **Négation** : ??/*Il n'est pas froid comme la glace.
- d. **Détachement** : ?/?/?Comme la glace, il est froid.
- e. **Actualisation** : Il est froid comme ?de la/??une /??cette/??/*Ø glace.
- f. **Métaphorisation** : ??(Cette personne) C'est la glace⁸⁸.
- g. **Substitution** : Il est ??refroidissant/*impassible comme la glace.

Contrairement à la locution précédente, la métaphore de glace donne ici plus de sens, même si elle semble peu naturelle. En effet, ceci est dû au fait que nous qualifions une personne et non une entité abstraite ou un geste exprimé par le pronom *ce*. La substitution synonymique pour un trait de comportement n'est pas correcte, car cela signifierait que la glace a le même trait que le comparé (?? la glace est impassible). D'ailleurs il est très

⁸⁶ La citation a été reprise telle qu'elle figurait dans le corpus.

⁸⁷ La date de publication et d'origine manque dans le corpus.

⁸⁸ Pour une personne qui donne l'apparence de l'insensibilité on dit plutôt « être un glaçon » (communication personnelle).

difficile de trouver un synonyme convenable de *froid* au sens propre du terme. En ce qui concerne la détermination relative aux articles, la locution avec *de la glace* est possible, car elle a été trouvée dans *FRANTEXT*, mais l'article partitif est peut-être contraint par le sens littéral du terme (cf. 3.2.3) :

(47) « *Le sol de terre battue, jonché de paille, de foin, de brindilles de bois était froid comme de la glace.* » (roman, Triolet, E., *Le premier accroc coûte 200 francs*, 1945)

D'où le caractère peut-être moins figé de cette locution que de l'autre.

En guise de conclusion, le nombre d'occurrences des deux locutions est très petit pour que nous puissions affirmer avoir trouvé de nouvelles locutions figées. Il faudrait avoir plusieurs sources différentes pour vérifier les hypothèses formulées. En réalité, le fait de déterminer une locution comme figée requiert une étude s'appuyant sur beaucoup plus de facteurs que le critère fréquentiel sur lequel nous étions basés. De plus, les évaluations des locuteurs natifs eux-mêmes sont parfois contradictoires, ce qui montre que la détermination des locutions comme figées n'est pas une tâche facile. Par ailleurs, le nombre réduit des occurrences peut être dû au fait que les expressions idiomatiques ne sont pas très fréquentes en général et que la plupart d'entre elles ne sont propres qu'à la langue orale. Notre démarche n'était donc qu'une brève initiation à cette problématique.

Conclusion

Le but de ce mémoire de licence était de présenter d'abord toutes les fonctions du mot *comme* et d'étudier par la suite son rôle dans une expression de comparaison.

Dans la partie théorique nous avons montré les tendances divergentes des dictionnaires et des grammaires pour classer ce mot dans une partie du discours. Pour nos analyses, nous avons considéré *comme* en tant que conjonction de subordination dans son emploi causal et temporel et en tant qu'adverbe de manière lorsque *comme* figure dans des phrases exclamatives ou comparatives. Concernant la comparaison, nous avons découvert que *comme* participe à l'expression de haut degré, et que cette dernière est souvent basée sur le procédé de métaphorisation. De plus, de telles comparaisons sont à l'origine de nombreuses locutions figées avec *comme*.

Nous nous sommes focalisés, dans la partie empirique, sur la construction « *être* + ADJ + *comme* », recherchée dans différents corpus linguistiques et ce en deux phases. En effet, l'hypothèse sur la présence de la construction dans n'importe quel type de discours est vraisemblable, mais avec une fréquence très faible dans tous les types de textes étudiés. De plus, l'hypothèse sur la distribution fréquentielle par rapport au type de texte a été confirmée. En effet, les textes littéraires emploient davantage la construction comparative avec *comme* et la comparaison métaphorique complexe est la plus marquante en poésie, comme présupposé. Par contre, les textes de spécialité reflètent une fréquence très faible de cette construction et la majorité des résultats y montrent une forte utilisation de la comparaison courante.

De plus, nous avons découvert que la suite la plus fréquente de *comme* est le SN et que ce dernier englobe le plus de comparaisons génériques de haut degré, prédominantes dans les textes poétiques, dramatiques et les sous-titres. Contrairement aux textes poétiques, nous avons détecté dans ces deux derniers types de textes beaucoup de comparants clichés.

Enfin, quatorze adjectifs parmi les plus fréquents ont été choisis comme candidats au figement. Nous avons découvert que ces adjectifs sont liés à des comparants très variés, dans des occurrences idiolectales des corpus comme dans des locutions déjà figées. Etant donné la grande diversité des occurrences, la recherche de nouvelles locutions figées s'est montrée problématique. Des locutions envisagées comme potentiellement figées ne

pouvaient pas être confirmées tout en prenant en considération le nombre d'occurrences peu suffisant et faute d'examen d'autres critères qui entrent en jeu.

D'une manière globale, le mémoire a contribué à analyser le mot *comme* dans son intégralité et il représente une initiation à la méthodologie de la linguistique de corpus.

En ce qui concerne les limites de la recherche, elles découlent principalement de la taille de certains sous-corpus et des faibles fréquences de la construction cherchée.

La composition textuelle des corpus joue un grand rôle aussi, car alors que *InterCorp* est très hétérogène et les textes littéraires y figurent en faible quantité, la majeure partie de *FRANTEXT* est formée de textes littéraires. De plus, la recherche ainsi que le traitement des occurrences obtenues dans ce dernier est assez incohérente et limitée. Effectivement, avec des outils informatiques plus performants, nous pourrions certainement obtenir des résultats plus pertinents. En outre, comme le codage grammatical n'est pas toujours correct, une grande partie des résultats étaient issus de bruits.

Les mesures associatives utilisées pour la recherche de nouvelles locutions figées ont été peu contributives à l'égard des occurrences de la construction très réduites. Sinon, l'identification automatique des locutions figées est souvent peu fiable, car très souvent, seulement l'introspection d'un locuteur natif pourrait en dévoiler plus sur le sujet.

Dans le futur, il serait convenable de compléter la recherche de la construction dans la nouvelle version d'*InterCorp* (version 8) plus élargie et en ajoutant des positions libres dans l'expression régulière. La recherche pourrait être complétée aussi par l'étude du voisinage dans *FRANTEXT* avec une application manuelle des mesures associatives.

Par ailleurs, il serait intéressant d'étudier les idiolectes des auteurs particuliers et peut-être consulter des locutions potentiellement figées auprès des locuteurs natifs sous forme d'un questionnaire. Une étude contrastive avec d'autres langues pourrait aussi apporter de nouvelles pistes pour la recherche.

Résumé en français

Ce mémoire de licence se donne pour but de s'interroger sur le mot *comme* et les emplois très variés de celui-ci en français contemporain, en se concentrant principalement sur l'analyse d'un emploi concret, à savoir la comparaison.

En ce qui concerne la partie théorique, elle est composée de trois chapitres. Dans le premier chapitre, après avoir consulté des dictionnaires et des grammaires de référence, nous faisons face à un classement problématique de *comme*, étant donné que ce dernier est rangé du point de vue fonctionnel parmi les adverbes ou les conjonctions voire les prépositions.

En effet, nous avons choisi comme emplois principaux de *comme* l'expression du rapport temporel et du rapport causal, l'exclamation et la comparaison. Nous avons fini par classer le *comme* temporel et causal avec une portée extra-prédicative dans le groupe des conjonctions et le *comme* exclamatif et comparatif avec une portée intra-prédicative parmi les adverbes. Ce classement va néanmoins à l'encontre de la théorie polysémique de certains linguistes, qui voient dans *comme* un pur marqueur de manière quel que soit son emploi.

D'ailleurs, nous avons également parcouru brièvement d'autres emplois plus ou moins marginaux, à savoir l'addition, l'attribution, l'approximation, l'exemplification et l'interrogation indirecte.

Dans le deuxième chapitre, nous nous penchons sur l'emploi temporel et causal de *comme* du point de vue syntaxique ainsi que sémantique. Nous avons appris que ces deux valeurs s'interpénètrent souvent. *Comme* figure ici en tant que conjonction de subordination dans des phrases circonstancielles et il est toujours le porteur du thème.

Le troisième chapitre se focalise sur le mot *comme* dans son statut d'adverbe de manière. Dans un premier temps, nous avons vu qu'il est à même d'exprimer la modalité intensive. Dans un second temps, ce qui est le plus pertinent pour nous, c'est l'usage de *comme* dans la comparaison. Au début nous présentons la typologie de cette dernière et nous nous familiarisons avec la terminologie appliquée à ce sujet. Quant à *comme*, il opère sur une comparaison qualitative de ressemblance et peut être suivi de différents contextes syntaxiques. Suivi d'un SN, le mot *comme* exprime « une identité de manière d'être », très souvent dans une structure elliptique. Il peut être suivi encore d'un SV exprimant dans ce

cas une « identité de manière de faire », d'un circonstant exprimant une « ressemblance entre situations » et bien d'autres.

Peut-être le trait le plus intéressant de *comme* est que celui-ci contribue, lié à un comparant parangon, à l'expression de haut degré couramment à la base d'une démarche métaphorique. En effet, beaucoup de comparaisons métaphoriques avec *comme* sont figées. La dernière partie de ce chapitre est consacrée à des notions telles que le de degré de figement et le défigement desquels nous profitons dans la partie empirique.

La partie empirique est réunie dans un chapitre et est destinée uniquement à une construction spécifique de *comme*, concrètement « être + ADJ + *comme* ».

En première phase, cette structure est cherchée d'abord dans le corpus *InterCop* dans lequel nous avons créé des sous-corpus selon les types de textes, à savoir les belles-lettres, les textes de spécialité, les sous-titres, les textes journalistiques, les textes juridiques et *Europarl*. Les autres corpus utilisés sont *L'Est républicain*, *frWaC* et *FRANTEXT*. Dans ce dernier ont été créés des sous-corpus de roman, de poésie, de théâtre et de textes de spécialité.

Nous avons confirmé notre hypothèse que la typologie textuelle joue un rôle important en matière fréquentielle de la construction. En effet, la construction est la plus fréquente dans des textes littéraires, notamment en poésie.

Pour ce qui touche l'analyse qualitative des occurrences de la construction, elle a dévoilé de nombreux bruits liés soit à des fautes dans le codage grammatical des corpus soit à des structures dans lesquelles *comme* exprime par exemple l'exclamation ou l'appartenance à une classe, et qui déviaient de notre construction comparative.

Nous avons procédé par la suite à l'étude de tous les contextes droits de *comme* et nous avons appris que la suite incontestablement la plus fréquente est le SN avec l'ellipse. De même, la révélation que le SN constitue le fondement pour des comparaisons génériques de haut degré nous conduit jusqu'à la deuxième phase où nous nous appuyons justement sur la structure « être + ADJ + *comme* + SN ». Dans cette section sont examinées des associations diverses pour comparer deux entités. D'un côté, il s'agit de comparaisons très complexes et originales dans des textes littéraires, en particulier dans la poésie, de l'autre côté nous avons trouvé beaucoup de comparants clichés, présents principalement dans les sous-titres, les textes dramatiques et les textes sur le web.

Les comparaisons clichées représentent l'objet de notre intérêt dans la dernière partie du travail. En nous servant des listes fréquentielles dans les corpus, nous avons sélectionné quatorze adjectifs candidats qui pourraient figurer dans des locutions figées avec *comme*. Ces dernières ont été cherchées auprès des dictionnaires pour pouvoir identifier des locutions figées dans les corpus.

En effet, nous avons découvert un large éventail de comparants pour des adjectifs choisis, que ce soit dans des locutions déjà figées ou dans des locutions idiolectales venant des corpus. La grande variété auprès des corpus nous a toutefois rendu difficile la recherche de nouvelles locutions figées. Dans l'intention d'en dénicher quelques-unes, nous avons utilisé des mesures associatives. Or, comme le nombre d'occurrences de la construction était peu important, les mesures n'étaient pas vraiment contributives. C'est pourquoi à la suite d'une analyse manuelle des résultats, nous avons choisi des locutions « *doux comme un mouton* », « *simple comme un coup de fil* » et « *froid comme la glace* » comme potentiellement figées. Il s'est montré finalement que la première est peut-être déjà figée dans le dictionnaire et que la deuxième est basée sur un slogan publicitaire. Des tests linguistiques appliqués aux deux dernières locutions n'étaient qu'une tentative pour déterminer leur degré de figement.

Pour conclure, le mémoire a contribué à résumer les connaissances sur le mot *comme* en français et à présenter plusieurs théories distinctes sur ce sujet. Le travail a aussi fait découvrir la méthodologie du travail sur les corpus linguistiques et même des limites de recherche dans ces derniers.

Résumé česky

Bakalářská práce si klade za cíl představit slovo *comme* a jeho různá užití v současné francouzštině s bližším zaměřením na jeho užití ve srovnáních.

Teoretická část je rozdělena do třech kapitol. V první kapitole narážíme na problematiku určení slovního druhu slova, jak ze strany referenčních slovníků, tak odborné literatury. *Comme* je totiž na základě konkrétní funkce chápáno jako spojka nebo příslovce, někdy dokonce jako předložka.

Za základní funkce slova bylo v této práci považováno vyjádření časového a kauzálního vztahu, zvolací modality a porovnání. Nakonec bylo *comme* posuzováno jako spojka s extrapredikační funkcí v prvních dvou a jakožto adverbium s funkcí intrapredikační v dalších dvou užitích. Tohle třídění se však nemusí shodovat s teoriemi některých lingvistů, opírajících se hlavně o polysémantický charakter slova, což znamená, že *comme* je ve všech svých užitích především indikátor způsobu.

V krátkosti jsou zmíněna i jiná, více či méně sekundární užití *comme*, a to slučovací vztah, doplňkové vyjádření příslušnosti k určité skupině, vyjádření neurčitosti, exemplifikace a také užití v nepřímých otázkách.

Druhá kapitola se věnuje podrobněji funkci slova ve větách časových a kauzálních ze syntaktického i sémantického hlediska. Zjistilo se, že se tyto dvě funkce běžně prolínají. *Comme* tu vystupuje jako hypotaktická spojka ve vedlejších větách příslovečných a je vždy v tematické pozici.

Ve třetí kapitole je *comme* nahlíženo ve svém postavení příslovce způsobu. V první řadě jsme viděli, že slovo dokáže vyjádřit intenzitu ve zvolacích větách.

Nejvíce podstatné je pro nás pak užití slova ve srovnávání. Nejdříve je definována typologie přirovnání jako taková a terminologie k ní se vztahují. Pokud jde o *comme*, to vyjadřuje míru podobnosti a může se pojít s různými syntaktickými kontexty. Následovně nominálním syntagmatem, *comme* vyjadřuje stejnou míru bytí a tahle struktura bývá často eliptická. Když předchází například slovesnému syntagmatu, odráží stejnou míru konání, nebo když předchází příslovečnému určení, odráží podobnost mezi dvěma situacemi.

Asi nejzajímavější vlastností *comme* je, že je spojeno s nekonkrétním srovnávaným pojmem, přispívá k vyjádření nejvyšší míry, často postavené na metaforickém základu. Mnoho

takových metaforických porovnání s *comme* jsou i ustálená. Poslední část této kapitoly definuje pojmy jako ustálenost a aktualizace, které jsou pak využívány v praktické části.

Praktická část představuje jednu rozsáhlou kapitolu věnovanou výlučně konstrukci « *être* + ADJ + *comme* ». V její první fázi se tato konstrukce vyhledává v korpusu *InterCorp*, kde byly vyčleněny subkorpusy beletrie, odborní texty, titulky, publicistické texty, právnické texty a texty ze zasedání Evropského parlamentu. Další korpusy, se kterými se pracuje, jsou *L'Est républicain*, *frWaC* a *FRANTEXT*. Ve *FRANTEXT*-u jsou ještě vymezeny subkorpusy podle literárních žánrů, konkrétně romány, poezie, divadelní texty a odborní texty.

Potvrdila se hypotéza, že frekvence hledané konstrukce je dána typem textu. Zjistilo se, že konstrukce je nejčastější v literatuře, zejména v poezii.

Co se týče kvalitativní analýzy jednotlivých výskytů konstrukce, bylo shledáno také mnoho šumů, které vyplynuli buďto z nesprávného gramatického značkování korpusu nebo z jiných než komparativních struktur, kde *comme* vyjadřovalo například zvolání nebo přiřazení k určité skupině.

Dále pak byly analyzovány všechny pravé kontexty slova *comme*, které byly nalezeny, a analýza jasně prokázala nejvyšší frekvenci pro nominální syntagma s elipsou. Zjištění, že tato struktura představuje základ pro komparace nejvyšší míry, vede pak k hlubší analýze konstrukce « *être* + ADJ + *comme* + SN » ve druhé fázi výzkumu. V této sekci se rozebírají různé možné asociace pro porovnání dvou entit. Na jedné straně jde o velmi originální a častokrát komplikovaná srovnání v literatuře na čele s poezií, na druhé straně se jedná o stereotypní spojení zastoupena především v titulcích a v dramatu.

Stereotypní spojení jsou předmětem zájmu poslední části téhle kapitoly. S využitím frekvenčních seznamů v korpusech bylo vybráno čtrnáct adjektiv jako kandidátů na tvorbu kolokací s *comme*. Nejdříve se konzultovali ve slovnících kolokace již existující, aby se teprve pak mohli identifikovat kolokace v korpusech.

Škála pro výběr kolokátů byla opravdu velká, jak pro spojení už lexikalizována, tak i pro ta idiolektická v korpusech. Velká variabilita kolokátů nám ovšem značně stěžila hledání nových kolokací. Pro jejich hledání jsme využili asociačních měr, aplikovaných v korpusech *KonText*-u. Nicméně vzhledem k tomu, že počet výskytů samotné konstrukce nebyl postačující, asociační míry nebyly moc přínosné. Proto byla potřebná i manuální

analýza výsledků, na základě které byly vybrány jako potenciálně ustálená následující spojení: « *doux comme un mouton* », « *simple comme un coup de fil* » a « *froid comme la glace* ». Nakonec se ukázalo, že první z nich už v slovníku figuruje a že druhá je postavena na reklamním sloganu. Lingvistické testy aplikované na dvě poslední spojení byly jenom pokusem o určení jejich míry ustálenosti.

Z celkového hlediska práce přispěla ke shrnutí poznatků slova *comme* ve francouzštině a představila různé přístupy k tomuto tématu. Byla rovněž ukázkou výzkumu v jazykových korpusech a jeho omezení.

Bibliographie

- BETH, A., & MARPEAU, E. (2005). *Figures de style*. S.l.: Librio.
- ČECHOVÁ, M., KRČMOVÁ, M., & MINÁŘOVÁ, E. (2008). *Současná stylistika* (1^e éd.). Prague: Nakladatelství Lidové noviny.
- ČERMÁK, F. (2007). *Frazeologie a idiomatika česká a obecná: Czech and general phraseology*. Prague: Karolinum.
- ČERMÁK, F. (2011). *Jazyk a jazykověda: přehled a slovníky* (4^e éd.). Prague: Karolinum.
- ČERMÁK, F., & ŠULC, M. (2006). *Kolokace* (1^e éd.). Prague: Nakladatelství Lidové noviny.
- DESMETS, M. (2008). Constructions comparatives en comme. *Langue Française*, 159(3), 33-49.
- DUBOIS, J. (c2012). *Dictionnaire de linguistique et des sciences du langage*. Paris: Larousse.
- FUCHS, C. (2014). *La comparaison et son expression en français*. Paris: Éditions Ophrys.
- FUCHS, C., & LE GOFFIC, P. (2008). Un emploi typifiant de « comme » : un de ces exemples comme on en trouve partout. *Langue Française*, 159(3), 67-82.
- GREVISSE, M., & GOOSSE, A. (c2008). *Le bon usage: grammaire française* : Grevisse langue française (14^e éd.). Bruxelles: Duculot.
- GROSS, G. (1996). *Les expressions figées en français: noms composés et autres locutions*. Paris: Ophrys.
- HENDRICH, J., RADINA, O., & TLASKAL, J. (2001). *Francouzská mluvnice* (3^e éd.). Plzeň: Fraus.
- KOKELBERG, J. (1993). *Les Techniques du style: vocabulaire, figures de rhétorique, syntaxe, rythme* (3^e éd.). Paris: Nathan.
- LE GOFFIC, P. (1994). *Grammaire de la phrase française*. Paris: Hachette.
- LOUCKÁ, H. (2010). *La communication verbale et le texte: Analyse linguistique de textes*. (1^e éd.). Prague: Karolinum.
- MOLINE, E. (2006). Belle comme un cœur / Belle comme sa sœur: Pour une description unifiée des comparatives en comme. *L'information Grammaticale*, 111(1), 14-20.
- MOLINE, E. (2008). Elle volait pour voler, comme on aime pour aimer : Les propositions d'analogie en comme. *Langue Française*, 159(3), 83-99.

- MOLINE, E., & FLAUX, N. (2008). Constructions en comme : homonymie ou polysémie ? Un état de la question. *Langue Française*, 159(3), 3-9.
- MONNERET, P., & RIOUL, R. (1999). *Questions de syntaxe française* (1^e éd.). Paris: PUF.
- NÁDVORNÍKOVÁ, O. (2012). *Korpusová analýza faktorů sémantické interpretace francouzského gérondivu*. Thèse. Filozofická fakulta Univerzity Karlovy v Praze. Dir. de thèse Loucká H.
- PECINA, P. (2009). *Lexical association measures: collocation extraction*. Prague: ÚFAL MFF UK.
- PORTINE, H. (1995). Fin comme (un) cheveu. *Revue De Linguistique Romane*, 59, 369-399.
- RIEGEL, M., PELLAT, J. -C., & RIOUL, R. (2014). *Grammaire méthodique du français* (5^e éd.). Paris: PUF.
- WILMET, M. (2007). *Grammaire critique du français*. Louvain-la-Neuve: Duculot.

Dictionnaires

Dictionnaire d'expressions idiomatiques [En ligne].

Disponible sur : <http://www.cnrtl.fr/dictionnaires/expressions_idiomatiques/>

Larousse : Dictionnaire de français [En ligne].

Disponible sur : <<http://www.larousse.fr/dictionnaires/francais>>

Le Nouveau Petit Robert Dictionnaire de la langue française 2008. (2007). [CD-ROM]. Paris: Dictionnaires Le Robert.

RAT, M. (1999). *Dictionnaire des expressions et locutions traditionnelles*. Paris : Larousse.

REY, A., & CHANTREAU, S. (1991). *Dictionnaire des expressions et locutions* (Nouv. éd.). Paris: Dictionnaires Le Robert.

Trésor de la langue française informatisé [En ligne].

Disponible sur : <<http://atilf.atilf.fr/tlf.htm>>

TLFi. Centre national de Ressources Textuelles et Lexicales [En ligne]

Disponible sur : <<http://www.cnrtl.fr/definition/>>

Corpus

FERRARESI, A. – BERNARDINI, S. – PICCI, G. – BARONI, M.: *frWaC*. Ústav Českého národního korpusu FF UK, Prague 2013. Disponible sur : <<http://www.korpus.cz>>

GAIFFE, B. – NEHBI, K.: *lEstRepublicain*. Ústav Českého národního korpusu FF UK, Prague 2013. Disponible sur : <<http://www.korpus.cz>>

FRANTEXT [en ligne]. Disponible sur : < <http://www.frantext.fr/>>

NÁDVORNÍKOVÁ, O. – VAVŘÍN, M.: *Korpus InterCorp* – français, version 7 de 19. 12. 2014. Ústav Českého národního korpusu FF UK, Prague 2014.

Disponible sur : <<http://www.korpus.cz>>

Liste des tableaux

- (1) **Tableau 1.1** : Quatre emplois principaux de *comme* mentionnés dans les dictionnaires, ceux-là divisés d'après le classement parmi les adverbes ou les conjonctions (✓ = mentionné/ ? = indéterminé/ gris = classé dans l'autre groupe) 11
- (2) **Tableau 1.2** : Quatre emplois principaux de *comme* mentionnés dans les grammaires, ceux-là divisés d'après le classement parmi les adverbes ou les conjonctions (✓ = mentionné/ ? = indéterminé/ gris = classé dans l'autre groupe) 13
- (3) **Tableau 4.1** : Taux de variabilité lexicale des corpus *InterCorp*, *L'Est républicain* et *frWaC* reçu du rapport types/tokens 34
- (4) **Tableau 4.2** : Fréquences absolues et pourcentages des occurrences analysées de la construction « être + ADJ + *comme* » dans tous les sous-corpus sélectionnés de *KonText* et de *FRANTEXT* après la suppression des « bruits » 44
- (5) **Tableau 4.3** : Répartition des occurrences analysées d'après le contexte droit de *comme* dans la construction « être + ADJ + *comme* » (après la suppression des bruits) dans les sous-corpus sélectionnés de *KonText* (excepté l'*Acquis communautaire* et *core – nonfiction*) et de *FRANTEXT* 45
- (6) **Tableau 4.4** : Fréquences absolues et pourcentages de comparaison figurative à SN dans la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *FRANTEXT* et de *KonText* (excepté l'*Acquis communautaire* et *core – nonfiction*) 50
- (7) **Tableau 4.5** : Fréquences absolues des associations comparatives dans chaque sous-corpus sélectionné de *KonText* et de *FRANTEXT* et fréquences absolues et pourcentages des associations comparatives dans tous les sous-corpus sélectionnés de *KonText* (excepté l'*Acquis communautaire* et *core – nonfiction*) et de *FRANTEXT* 54
- (8) **Tableau 4.6** : Dix premiers rangs dans les listes fréquentielles des corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC* avec les adjectifs candidats au figement marqués en couleurs différentes et les fréquences absolues de ces derniers 57
- (9) **Tableau 4.7** : Fréquences absolues et pourcentages des adjectifs candidats au figement dans la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *KonText* (excepté l'*Acquis communautaire*) 58
- (10) **Tableau 4.8** : Fréquences absolues et pourcentages des adjectifs candidats au figement dans la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *FRANTEXT* 58
- (11) **Tableau 4.9** : Fréquences relatives (ipm) des adjectifs candidats au figement hors la construction « être + ADJ + *comme* » dans les corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC* 59
- (12) **Tableau 4.10** : Mots entrant en locutions figées avec *comme* et les adjectifs candidats au figement dans la construction « être + ADJ + *comme* », trouvés dans les dictionnaires utilisés; mots trouvés aussi dans les corpus sont marqués en gras 61
- (13) **Tableau 4.11** : Mots entrant en locution (autre que figée) avec *comme* et les adjectifs candidats au figement dans la construction « être + ADJ + *comme* » dans les corpus sélectionnés de *KonText* (*InterCorp* excepté l'*Acquis communautaire*) et de *FRANTEXT* 63

- (14) **Tableau 4.12** : Lemmes entrant en collocation avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans les corpus *InterCorp* (excepté l'*Acquis communautaire*), *L'Est républicain* et *frWaC* ; avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*..... 65
- (15) **Tableau 4.13** : Locutions avec *comme* et les adjectifs candidats au figement, à savoir *froid*, *simple* et *doux*, issues de la construction « être + ADJ + *comme* » dans les corpus sélectionnés de *KonText* (*InterCorp* excepté l'*Acquis communautaire*) et de *FRANTEXT*, choisies comme potentiellement figées ; avec l'affichage du n° de rang dans la liste de collocations, de la valeur de la mesure *MI* et de la fréquence absolue dans le corpus..... 66

Liste des graphiques

- (16) **Graphique 4.1** : Fréquence relative (ipm) de la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *KonText*, avec les fréquences absolues indiquées à l'intérieur du graphique 35
- (17) **Graphique 4.2** : Taux de variabilité lexicale de la construction « être + ADJ + *comme* » en pourcents dans les sous-corpus sélectionnés de *KonText* (excepté *core – nonfiction*), reçu du rapport types/tokens..... 36
- (18) **Graphique 4.3** : Fréquence relative (ipm) de la construction « être + ADJ + *comme* » dans les sous-corpus sélectionnés de *FRANTEXT*, avec les fréquences absolues indiquées à l'intérieur du graphique..... 39
- (19) **Graphique 4.4** : Comparaison des fréquences relatives (ipm) de la construction « être + ADJ + *comme* » dans tous les sous-corpus sélectionnés de *KonText* et de *FRANTEXT* (excepté *core – nonfiction*) 40

ANNEXE A

CONJONCTION	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Le Petit Robert	?	?		0	?
Larousse	✓	✓	0	✓	✓
TLFi	?	0	0	?	0
ADVERBE	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Le Petit Robert	?	?	✓	0	?
Larousse			0		
TLFi	?	0	0	?	0

Tableau A.1 : Emplois secondaires de *comme* mentionnés dans les dictionnaires divisés d'après le classement parmi les adverbess ou les conjonctions (✓ = mentionné/ 0= absent/ ?= indéterminé/ gris= classé dans l'autre groupe)

CONJONCTION	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Fr. mlvnicce	0	✓	0	0	0
G. de la Phrase Fr.			0	0	0
Le bon usage	0	✓		0	?
Grammaire méth.	0	0	0	0	0
ADVERBE	Addition	Attribution	Interrog. indirecte	Exemplification	Approximation
Fr. mlvnicce	0		0	0	0
G. de la Phrase Fr.	✓	✓	0	0	0
Le bon usage	0		✓	0	?
Grammaire méth.	0	0	0	0	0

Tableau A.2 : Emplois secondaires de *comme* mentionnés dans les grammaires divisés d'après le classement parmi les adverbess ou les conjonctions (✓ = mentionné/ 0= absent/ ?= indéterminé/ gris= classé dans l'autre groupe)

ANNEXE B

sous-corpus <i>KonText</i>	N° de tokens
frWaC	1 613 814 684
L'Est républicain	146 544 420
sous-titres	32 897 232
Acquis communautaire	32 523 605
Europarl	18 968 039
core – fiction	7 194 038
PressEurop, Syndicate	6 311 555
core – nonfiction	816 023

Tableau B.1 : Taille en nombre de positions (tokens) des sous-corpus d'*InterCorp* et des corpus *L'Est républicain* et *frWaC*

sous-corpus <i>KonText</i>	f absolue	f relative	N° de types
frWaC	12 475	7,73	1740
sous-titres	735	22,34	236
L'Est républicain	549	3,75	228
core – fiction	203	28,22	120
Acquis communautaire	66	2,03	19
Europarl	28	1,48	24
PressEurop, Syndicate	20	3,17	17
core – nonfiction	1	1,23	1

Tableau B.2 : Fréquences absolues et relatives (ipm) et les nombres de types de lemmes de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *KonText*; sous-corpus rangés dans l'ordre décroissant selon la fréquence absolue

sous-corpus <i>FRANTEXT</i>	f absolue	f relative
roman 1900 – 1997	1439	38,82
roman 1920 – 1939	451	38,71
traités, essais 1900 – 1997	287	12,37
roman 1960 – 1979	291	39,98
roman 1940 – 1959	264	37,54
roman 1980 – 1997	230	38,39
théâtre 1900 – 1997	229	66,82
roman 1900 – 1919	203	39,67
poésie 1900 – 1997	196	101,09

Tableau B.3 : Fréquences absolues et relatives de la construction « être + ADJ + comme » dans les sous-corpus sélectionnés de *FRANTEXT* rangés dans l'ordre décroissant selon la fréquence absolue

sous-corpus <i>FRANTEXT</i>	N° de textes
roman 1900 – 1919	68
roman 1920 – 1939	164
roman 1940 – 1959	82
roman 1960 – 1979	80
roman 1980 – 1997	80
poésie 1900 – 1997	111
théâtre 1900 – 1997	127
traités, essais 1900 – 1997	399

Tableau B.4 : Nombres de textes trouvés dans *FRANTEXT* pour les sous-corpus sélectionnés

ANNEXE C

rang	InterCorp logDice	f	T- score	MI	logDice	rang	InterCorp MI	f	T- score	MI	logDice
1.	roc	8	2.828	12.542	7.866	1.	barrique	4	2.000	13.750	7.007
2.	agneau	10	3.160	10.795	7.722	2.	pinson	4	2.000	13.750	7.007
3.	neige	16	3.996	9.871	7.603	3.	gardon	3	1.732	12.801	6.574
4.	linge	10	3.160	10.485	7.602	4.	berceuse	3	1.732	12.669	6.568
5.	carpe	5	2.235	11.529	7.143	5.	roc	8	2.828	12.542	7.866
6.	mule	5	2.235	10.976	7.048	6.	carpe	5	2.235	11.529	7.143
7.	barrique	4	2.000	13.750	7.007	7.	mule	5	2.235	10.976	7.048
8.	pinson	4	2.000	13.750	7.007	8.	braise	3	1.731	10.853	6.415
9.	chaton	4	1.999	10.660	6.727	9.	agneau	10	3.160	10.795	7.722
10.	gardon	3	1.732	12.801	6.574	10.	chaton	4	1.999	10.660	6.727
11.	berceuse	3	1.732	12.669	6.568	11.	linge	10	3.160	10.485	7.602
12.	plume	7	2.640	8.801	6.487	12.	pou	3	1.730	10.053	6.272
13.	pot	6	2.445	8.943	6.476	13.	neige	16	3.996	9.871	7.603
14.	clou	4	1.997	9.547	6.426	14.	chêne	3	1.730	9.867	6.229
15.	braise	3	1.731	10.853	6.415	15.	clou	4	1.997	9.547	6.426
16.	cochon	6	2.443	8.663	6.315	16.	pot	6	2.445	8.943	6.476
17.	pou	3	1.730	10.053	6.272	17.	balai	3	1.728	8.801	5.886
18.	chêne	3	1.730	9.867	6.229	18.	plume	7	2.640	8.801	6.487
19.	miel	4	1.994	8.361	5.893	19.	cochon	6	2.443	8.663	6.315
20.	balai	3	1.728	8.801	5.886	20.	miel	4	1.994	8.361	5.893
21.	glace	6	2.440	7.953	5.850	21.	planche	3	1.727	8.347	5.686
22.	enfer	6	2.439	7.865	5.787	22.	glace	6	2.440	7.953	5.850
23.	soleil	9	2.985	7.610	5.721	23.	enfer	6	2.439	7.865	5.787
24.	planche	3	1.727	8.347	5.686	24.	soleil	9	2.985	7.610	5.721
25.	chien	11	3.293	7.113	5.335	25.	chien	11	3.293	7.113	5.335
26.	bonjour	13	3.570	6.664	4.949	26.	pierre	4	1.982	6.777	4.822
27.	pierre	4	1.982	6.777	4.822	27.	bonjour	13	3.570	6.664	4.949
28.	dieu	3	1.714	6.623	4.608	28.	dieu	3	1.714	6.623	4.608
29.	roi	4	1.973	6.192	4.344	29.	ca	3	1.709	6.217	4.292
30.	ressemble r	4	1.972	6.136	4.296	30.	roi	4	1.973	6.192	4.344

31.	ca	3	1.709	6.217	4.292	31.	ressemble r	4	1.972	6.136	4.296
32.	beurre	3	1.704	5.968	4.089	32.	beurre	3	1.704	5.968	4.089
33.	air	9	2.940	5.648	3.962	33.	air	9	2.940	5.648	3.962
34.	pied	5	2.189	5.568	3.830	34.	pied	5	2.189	5.568	3.830
35.	boulot	3	1.694	5.520	3.709	35.	boulot	3	1.694	5.520	3.709
36.	bras	3	1.692	5.447	3.645	36.	bras	3	1.692	5.447	3.645
37.	mort	8	2.755	5.261	3.586	37.	client	3	1.690	5.362	3.570
38.	client	3	1.690	5.362	3.570	38.	cœur	5	2.179	5.284	3.569
39.	cœur	5	2.179	5.284	3.569	39.	mort	8	2.755	5.261	3.586

Tableau C.1 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *InterCorp* (excepté *l'Acquis communautaire*), avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*

rang	L'Est rép. logDice	f	T- score	MI	logDice	rang	L'Est rép. MI	f	T- score	MI	logDice
1.	gardon	9	2.999	12.277	8.157	1.	gardon	9	2.999	12.277	8.157
2.	sou	5	2.235	10.772	6.964	2.	sou	5	2.235	10.772	6.964
3.	œuf	10	3.158	9.535	6.304	3.	gosse	3	1.730	10.099	6.264
4.	gosse	3	1.730	10.099	6.264	4.	œuf	10	3.158	9.535	6.304
5.	orange	3	1.728	8.773	5.368	5.	orange	3	1.728	8.773	5.368
6.	poisson	8	2.820	8.364	5.220	6.	poisson	8	2.820	8.364	5.220
7.	pomme	4	1.990	7.623	4.458	7.	pomme	4	1.990	7.623	4.458
8.	témoigner	4	1.989	7.531	4.374	8.	Il	3	1.723	7.541	4.341
9.	Il	3	1.723	7.541	4.341	9.	témoigner	4	1.989	7.531	4.374
10.	image	8	2.805	6.937	3.866	10.	image	8	2.805	6.937	3.866
11.	pierre	3	1.716	6.728	3.600	11.	pierre	3	1.716	6.728	3.600
12.	neige	3	1.714	6.617	3.497	12.	neige	3	1.714	6.617	3.497
13.	habitude	3	1.714	6.597	3.478	13.	habitude	3	1.714	6.597	3.478
14.	Jean-Paul	3	1.713	6.531	3.416	14.	Jean-Paul	3	1.713	6.531	3.416
15.	coup	14	3.698	6.417	3.373	15.	coup	14	3.698	6.417	3.373
16.	remplacer	3	1.711	6.367	3.261	16.	remplacer	3	1.711	6.367	3.261
17.	suivre	10	3.100	5.661	2.620	17.	suivre	10	3.100	5.661	2.620
18.	gens	4	1.960	5.637	2.575	18.	gens	4	1.960	5.637	2.575
19.	cela	13	3.527	5.518	2.481	19.	cela	13	3.527	5.518	2.481
20.	monde	8	2.754	5.245	2.205	20.	monde	8	2.754	5.245	2.205

21.	la le	16	3.875	4.999	1.967	21.	la le	16	3.875	4.999	1.967
22.	celui	12	3.341	4.810	1.777	22.	celui	12	3.341	4.810	1.777
23.	chaque	6	2.358	4.749	1.709	23.	chaque	6	2.358	4.749	1.709
24.	!	6	2.290	3.938	0.904	24.	!	6	2.290	3.938	0.904
25.	dire	4	1.866	3.903	0.866	25.	dire	4	1.866	3.903	0.866
26.	un	119	10.139	3.824	0.798	26.	un	119	10.139	3.824	0.798
27.	jeu	3	1.592	3.626	0.588	27.	jeu	3	1.592	3.626	0.588
28.	trois	5	2.050	3.583	0.550	28.	trois	5	2.050	3.583	0.550
29.	on	10	2.895	3.565	0.535	29.	on	10	2.895	3.565	0.535
30.	jour	6	2.236	3.519	0.488	30.	jour	6	2.236	3.519	0.488

Tableau C.2 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *L'Est républicain*, avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*

rang	frWaC logDice	f	T- score	MI	logDice	rang	frWaC MI	f	T- score	MI	logDice
1.	orange	27 9	16.689	10.162	7.700	1.	bbbbrrrrrrrrrr	3	1.732	16.981	2.978
2.	agneau	34	5.818	8.797	5.639	2.	repos.de	5	2.236	16.718	3.714
3.	rétractilité	17	4.123	16.609	5.478	3.	rétractilité	17	4.123	16.609	5.478
4.	braise	20	4.468	10.139	5.472	4.	porçin	3	1.732	16.566	2.978
5.	chou	32	5.636	8.086	5.242	5.	drv	4	2.000	14.337	3.390
6.	hab	20	4.459	8.435	5.038	6.	botin	4	2.000	13.981	3.390
7.	neige	61	7.762	7.343	5.032	7.	bronchitique	3	1.732	13.866	2.975
8.	roc	15	3.867	9.447	5.010	8.	tulipa	3	1.732	13.708	2.975
9.	carpe	17	4.114	8.785	4.996	9.	Obligé	4	2.000	13.148	3.387
10.	garde-fou	13	3.603	10.265	4.944	10.	calzone	3	1.732	12.981	2.973
11.	pou	15	3.863	8.570	4.805	11.	soue	3	1.732	12.357	2.970
12.	sensation	40	6.281	7.185	4.770	12.	Beaudelaire	5	2.236	12.314	3.701
13.	habitude	60	7.677	6.809	4.589	13.	Crésus	7	2.645	12.189	4.178
14.	cochon	18	4.220	7.541	4.561	14.	coruscant	3	1.732	11.880	2.966
15.	mule	10	3.157	9.199	4.481	15.	limande	6	2.449	11.500	3.948
16.	Boîte	11	3.307	8.368	4.425	16.	chiffe	3	1.731	11.500	2.963
17.	éclair	14	3.722	7.552	4.374	17.	sylvestris	3	1.731	11.337	2.961
18.	taupe	9	2.995	9.181	4.348	18.	claustra	3	1.731	11.337	2.961
19.	bonjour	74	8.507	6.489	4.348	19.	pinson	7	2.644	10.581	4.134

20.	balai	10	3.151	8.077	4.246	20.	temperatures	3	1.731	10.538	2.948
21.	éternité	13	3.582	7.286	4.196	21.	democrates	3	1.731	10.432	2.946
22.	oeuf	44	6.556	6.424	4.194	22.	popups	3	1.731	10.328	2.944
23.	Crésus	7	2.645	12.189	4.178	23.	Ritchie	3	1.731	10.295	2.943
24.	pinson	7	2.644	10.581	4.134	24.	garde-fou	13	3.603	10.265	4.944
25.	kite	7	2.643	9.898	4.095	25.	orange	27 9	16.689	10.162	7.700
26.	témoigner	38	6.087	6.306	4.061	26.	braise	20	4.468	10.139	5.472
27.	dieu	28	5.228	6.380	4.035	27.	kite	7	2.643	9.898	4.095
28.	truc	42	6.394	6.230	4.019	28.	Crochet	4	1.998	9.731	3.324
29.	cul	12	3.437	6.977	3.987	29.	pivoine	4	1.998	9.673	3.322
30.	linge	14	3.707	6.742	3.964	30.	Sevan	3	1.730	9.618	2.922
31.	limande	6	2.449	11.500	3.948	31.	roc	15	3.867	9.447	5.010
32.	tigre	8	2.813	7.538	3.861	32.	Hérode	6	2.446	9.300	3.842
33.	Hérode	6	2.446	9.300	3.842	33.	mule	10	3.157	9.199	4.481
34.	tomate	20	4.413	6.240	3.811	34.	taupe	9	2.995	9.181	4.348
35.	planter	21	4.521	6.209	3.806	35.	Winehouse	4	1.996	9.118	3.289
36.	ornement al	6	2.443	8.508	3.751	36.	gardon	5	2.232	9.083	3.583
37.	repos.de	5	2.236	16.718	3.714	37.	pastèque	3	1.728	8.910	2.888
38.	Beaudelair e	5	2.236	12.314	3.701	38.	teigne	3	1.728	8.880	2.886
39.	papillon	10	3.131	6.655	3.694	39.	cole	3	1.728	8.833	2.883
40.	âne	9	2.973	6.811	3.690	40.	agneau	34	5.818	8.797	5.639
41.	rapace	6	2.440	7.959	3.656	41.	carpe	17	4.114	8.785	4.996
42.	amortisse ur	6	2.438	7.758	3.614	42.	ultralibéral	4	1.995	8.638	3.250
43.	camion	13	3.558	6.233	3.613	43.	pou	15	3.863	8.570	4.805
44.	mec	14	3.689	6.141	3.589	44.	fondeur	4	1.995	8.567	3.244
45.	mouchoir	6	2.437	7.628	3.584	45.	ornemental	6	2.443	8.508	3.751
46.	gardon	5	2.232	9.083	3.583	46.	rj	3	1.727	8.500	2.859
47.	singe	8	2.803	6.779	3.581	47.	hab	20	4.459	8.435	5.038
48.	ange	16	3.936	5.959	3.519	48.	linceul	3	1.727	8.410	2.852
49.	sou	8	2.798	6.538	3.473	49.	Boîte	11	3.307	8.368	4.425
50.	vôtre	8	2.797	6.513	3.461	50.	cerf-volant	4	1.994	8.266	3.211

Tableau C.3 : Lemmes entrant en locutions avec *comme* et un adjectif dans la construction « être + ADJ + *comme* » dans le corpus *frWaC*, avec les valeurs des mesures associatives *MI*, *T-score* et *logDice* affichées et les rangs ordonnés d'après *MI* ou *logDice*