

Les degrés d'intersection de l'équivalence sémantique approchée en parataxe de mots

Gaëlle Doualan

Université Paris-Sorbonne
Sens, Texte, Histoire, Informatique (EA 4509)
gaelle.doualan@gmail.com

Résumé. L'étude historique et épistémologique de la synonymie montre l'importance du recentrage de la notion sur les équivalences sémantiques qui la fonde, ce qui suppose de la penser dans une démarche onomasiologique. Dans sa perspective différentialiste, l'étude de la synonymie s'appuie sur le test de la substitution des mots en contexte mais si ce test est recevable dans une linguistique qui prend la phrase comme unité de base, il ne l'est plus dans une linguistique du texte. Les synonymes sont donc étudiés selon leur répartition dans le texte, ce qui revient à rechercher des réseaux lexicaux et donc à étudier des isotopies. Cette étude nécessite une large couverture lexicale et une structuration du lexique étudié apportées par la démarche onomasiologique. L'étude textuelle et onomasiologique de la synonymie est menée à partir du vocabulaire du vice et de la vertu sur un corpus de trois oeuvres de la littérature morale du XVIIIe siècle. Le corpus est segmenté en paragraphes et une ressource lexicale structurant le vocabulaire du vice et de la vertu selon les trois principales relations sémantiques que sont la synonymie, l'antonymie et l'hyponymie est créée. Le corpus et le thesaurus étant au format XML, un traitement automatique est mis en place à l'aide d'une requête XQuery : le thesaurus est projeté sur le corpus ce qui permet de révéler les réseaux lexicaux dans les paragraphes du corpus. L'analyse linguistique menée à partir des résultats obtenus consiste à approfondir l'étude de la synonymie en s'intéressant particulièrement à l'équivalence sémantique telle qu'elle se manifeste en contexte. L'équivalence sémantique est décomposée en trois degrés d'intersection : une intersection forte qui correspond à une interdéfinition entre synonymes, une intersection intermédiaire qui manifeste la créativité linguistique au travers des rapprochements produits par les locuteurs en contexte et une intersection faible qui correspond à la cohyponymie et qui se résume à la possession en commun d'un seul sème (micro)générique. L'étude des équivalences sémantiques en contexte révèle les différences entre langue et discours : si en discours, on retrouve des relations établies en langue, soient les intersections forte et faible, l'intersection intermédiaire est l'espace de liberté du discours par rapport à la langue et constitue une manifestation de la déformabilité et de l'idiosyncrasie dans la langue.

Abstract. The historical and epistemological study of synonymy has shown the importance of focusing on the notion of semantic equivalences because of how they shape synonymy. This means it is necessary to consider the synonymy from an onomasiological approach. From a differentialist point of view, the study of synonymy relies on the substitution of words in context, but if substitution is possible in sentence-based linguistics, it is not any more possible in a text-based linguistics.

Therefore synonyms are studied in texts, focusing on their distribution, what amounts to look for lexical networks and to study isotopies. This onomasiological study needs a great number of words structured in a thesaurus. It is managed with the vocabulary of vice and virtue on a corpus of three texts belonging to seventeenth century French moral literature. Each text is tokenized in paragraphs and a lexical resource is created: it structures the vocabulary of vice and virtue with the three main semantic relations, in other words, synonymy, antonymy and hypernymy. Both corpus and thesaurus are built in XML format, allowing an automatic processing with a XQuery request. The thesaurus is applied to the corpus to find lexical networks in the paragraphs of the corpus. A linguistic analysis is conducted on these results to delve deeper into the study of synonymy and especially the contextual semantic equivalences which are split up into three degrees of semantic intersection: a high intersection that corresponds to synonyms which are interdefined, a medium intersection that shows the linguistic creativity of the writer through bringing words closer contextually and a low intersection that corresponds to cohyponymy and relies on only one common seme. The study of semantic equivalences in context reveals the differences between language and discourse: if the high and low intersections are common to both discourse and language, the medium intersection is the space of freedom in the discourse in comparison with the language and it constitutes the expression of deformability and idiosyncrasy in discourse.

1 Introduction

L'étude de l'histoire de la synonymie (Doualan, 2015) révèle que cette notion, onomasiologique par définition¹, est traitée selon une démarche sémasiologique dans la conception différentialiste ; la tradition synonymiste² puis les sémantiques structuraliste³ et cognitive⁴ s'attachent à rechercher les différences de sens qui résident entre les mots synonymes. Or la synonymie, en tant d'équivalence de sens, est la possibilité de répertorier plusieurs instanciations lexicales pour un même signifié ; elle s'affilie donc à l'onomasiologie en tant qu'étude des signifiants se rapportant à un même signifié. A l'inverse, la recherche des signifiés d'un ou plusieurs signes ressortit d'une démarche sémasiologique et se rapporte à l'étude de la polysémie. Aussi, réorienter l'étude de la synonymie vers l'onomasiologie est un retour à ce qui lui est consubstantielle et la sépare de toute étude distinctive pour se focaliser sur la communauté de sens. L'étude onomasiologique de la synonymie a pour objectif de décomposer la communauté de sens⁵ et de comprendre le mécanisme de rapprochement lexical : comment décèle-t-on des synonymes ? Sur quoi repose l'équivalence sémantique entre les synonymes ? Cette équivalence est-elle modulable ?

Le sens étant immanent au texte, la synonymie, en tant que relation sémantique, est étudiée en discours et non en langue – comme c'est le cas chez les synonymistes. Lorsqu'elle est restreinte à la langue, la synonymie n'est qu'un artefact de la lexicographie, alors qu'elle a tout lieu d'être étudiée en discours, puisqu'elle participe à la construction dynamique du sens dans les textes. La synonymie est étudiée en contexte par les sémanticiens grâce au test de la substitution. Ce test modifie l'observable : on substitue un mot présent dans un contexte à l'un de ses synonymes pour tester leurs sens. Ce test est recevable dans une linguistique de la phrase mais, dans une linguistique du texte, il ne tient pas compte du processus rédactionnel inhérent à tout texte⁶. Aussi, dès lors que l'on travaille dans une perspective textuelle, il faut proposer une méthode d'analyse qui ne modifie pas l'observable. Pour l'étude de la synonymie, on suggère de révéler et d'analyser la répartition des synonymes dans un texte. Cela revient à rechercher les réseaux lexicaux ou isotopies qui traversent le texte. Cette méthode nécessite une grande quantité d'unités lexicales de sens proches, mais qui n'appartiennent pas forcément à la même catégorie grammaticale⁷. La perspective discursive se conjugue donc à l'onomasiologie pour renouveler l'étude de la synonymie. Des comparaisons peuvent être effectuées entre les relations de synonymie rencontrées en discours et celles qui sont répertoriées dans les dictionnaires afin de mesurer la déformabilité de la langue (Culioli, 1990) eu égard aux relations sémantiques.

L'équivalence sémantique approchée qui fonde la synonymie est étudiée en contexte dans les réseaux lexicaux qu'elle traverse. Les équivalences relevées sont classées selon trois degrés d'intersection

sémantique – forte, intermédiaire et faible – qui rendent compte du caractère modulable du rapprochement entre les mots. Cette décomposition s’appuie sur le nombre de sèmes communs aux unités lexicales reliées par l’équivalence. Les intersections fortes et faibles correspondent à des relations sémantiques répertoriées en langue, alors que l’intersection intermédiaire témoigne de la créativité des locuteurs en matière de relations sémantiques et donc de la déformabilité de la langue. Pour neutraliser l’influence sémantique du contexte sur les relations d’équivalences, les réseaux lexicaux sont recherchés dans de faibles emfans textuels. Mais une fois une classification des réseaux lexicaux obtenue, il sera possible d’étudier ces réseaux et donc les relations sémantiques dans des emfans textuels de toute taille.

Les hypothèses théoriques et le protocole expérimental qui ont présidé à l’étude onomasiologique de la synonymie sont décrits dans les deux premières parties. La troisième partie rend compte parmi les réseaux lexicaux obtenus expérimentalement de ceux qui témoignent d’équivalences approchées ; ils sont classés selon les trois degrés d’intersection sémantique et selon les thèmes choisis.

2 Étude onomasiologique et discursive de la synonymie

L’étude onomasiologique et discursive de la synonymie s’appuie sur la méthode développée en analyse thématique. Les spécificités de cette analyse sont présentées ; puis sont exposées les modifications qu’elle a subies afin de se conformer à l’étude de la synonymie.

2.1 Analyse thématique

La thématique, composante de la sémantique interprétative (Rastier, 1987), « rend compte des contenus investis, c’est-à-dire du secteur de l’univers sémantique mis en œuvre dans le texte. Elle en décrit les unités » (Rastier *et al.*, 1994 : 40). Elle se place du côté du signifié – le contenu – mais s’intéresse aussi aux unités de ce contenu, ce qui concerne aussi bien les traits sémantiques que les signes qui comportent ces traits sémantiques et *a fortiori* leurs signifiants. La thématique décrit les thèmes abordés dans un texte et étudie les instanciations lexicales de ces thèmes, c’est-à-dire, leurs signifiants, appelés lexicalisations en analyse thématique, puisqu’elles sont des mises en lexique des thèmes. La méthodologie de l’analyse thématique dépend de l’objectif visé (Martin, 1995 : 18), d’où le caractère subjectif de cette étude au point qu’il existe une grande diversité d’études thématiques (cf. Rastier, 1995). Ces études dépendent du choix du corpus : quel type de textes ? Quel genre ? Le corpus est-il constitué de textes d’un même auteur ou d’auteurs différents (Martin, 1995 : 18) ? Les paramètres de choix du corpus influencent le choix du thème qui préside à l’étude mais aussi le choix de la méthode qui est mise en place pour l’étude du thème. Par ailleurs, il est possible d’étudier un ou plusieurs thèmes, voire de rechercher les thèmes connexes à un thème donné dans un corpus.

2.1.1 Le thème

L’analyse thématique repose sur la notion de thème qui se définit comme « une structure stable de traits sémantiques (ou *sèmes*), récurrente dans un corpus, et susceptible de lexicalisations diverses » (Rastier, 2001 : 197). Le thème rassemble les paramètres nécessaires à une étude onomasiologique et discursive de la synonymie, puisqu’il est un signifié, c’est-à-dire, un ensemble de sèmes, il est à rechercher en discours, et il se manifeste selon divers signifiants. Dans une analyse thématique, le thème sélectionné « ne peut avoir qu’une définition et une seule, et celle-ci doit être connue dès l’abord » (Martin, 1995 : 16). Le flou de la notion ne doit pas empêcher de définir de manière précise le ou les thèmes abordés dans une étude. Pour la manifestation du thème, une grande latitude existe également car plusieurs types de signifiants sont possibles : « le thème est représenté par une séquence linguistique (une phrase, un groupe nominal, un nom propre ou commun) » (Erlich, 1995 : 85).

Toutefois, le thème ne se réduit pas à un lexème, il n’est pas un signe, même s’il peut être directement manifesté par une lexicalisation qui prend le statut de lexicalisation privilégiée du thème. Si le thème était un lexème, l’analyse thématique se confondrait avec la lexicographie et la lexicologie et perdrait son

intérêt. Le thème est une construction qui dépend du choix du corpus, des choix méthodologiques et de l'objet d'étude visé. Ainsi, la définition d'un thème se fait toujours en indiquant les traits sémantiques qui le composent, les lexicalisations qui le manifestent et le corpus dans lequel il est étudié. Tous ces paramètres sont constitutifs du thème et le modèle de telle sorte qu'ils doivent être systématiquement précisés.

2.1.2 Les lexicalisations du thème

Les lexicalisations sont des instanciations lexicales qui manifestent un thème dans la mesure où elles comportent tout ou partie des sèmes constituant ce thème. « Un thème, ainsi défini comme molécule sémique, peut ainsi recevoir des expressions diverses, par des unités qui vont du morphème au syntagme. Nous les nommerons, pour simplifier, *lexicalisations* » (Rastier, 2001 : 200). Comme il n'est pas possible d'être exhaustif dans l'inventaire des lexicalisations d'un thème, il faut fixer des limites au choix des lexicalisations⁸. Si les lexicalisations sont choisies *a priori*, elles sont confrontées à la réalité du corpus avant d'être retenues pour l'analyse thématique. Le corpus « n'est accessible que par extraits interrogeables par *mots-clés*. Il nous est donc nécessaire de traduire le thème en une liste de mots-clés que nous appellerons *formes d'entrée* » (Erllich, 1995 : 85). D'un point de vue pratique, les lexicalisations – ou formes d'entrée – sont sélectionnées si elles peuvent « permettre la sélection d'un grand nombre d'extraits qui illustrent le thème (*Ibid.*). Plusieurs lexicalisations du thème entretenant des relations sémantiques dans un extrait du corpus forment un réseau lexical.

Si la lexicalisation privilégiée est le plus souvent un substantif, les autres catégories grammaticales fournissent des lexèmes pertinents pour l'étude d'un thème. La dérivation et le recours aux synonymes enrichissent l'inventaire des lexicalisations et accroissent la couverture du thème. L'étude d'un thème peut être étendue aux cooccurrents des lexicalisations, qui prennent le statut de corrélats, s'ils sont sémantiquement apparentés au thème. Cela permet de saisir les thèmes connexes au thème étudié et de décrire les réseaux associatifs entre lexicalisations et corrélats et *a fortiori* les interactions entre les thèmes au sein du corpus.

2.2 L'analyse thématique appliquée à la synonymie

2.2.1 De nombreuses lexicalisations et un thème sémantiquement large

L'analyse thématique est une étude textuelle qui devient onomasiologique lorsqu'elle est centrée sur les lexicalisations et même sur un grand nombre de lexicalisations. Dans l'étude de la synonymie, l'accent est mis sur les lexicalisations et sur les réseaux lexicaux qu'elles forment plus que sur le thème en lui-même. Une centaine de lexicalisations est retenue pour la présente étude. Un thème sémantiquement large est nécessaire pour susciter une telle quantité de lexicalisations. La collecte des lexicalisations se fait par enrichissements successifs. À partir d'une lexicalisation privilégiée, les dérivés morphologiques sont ajoutés puis des synonymes, des hyponymes et leurs dérivés. Les hyponymes sont eux-mêmes l'objet d'un enrichissement à l'aide de synonymes. Cette récolte se fait en langue mais elle est validée par une confrontation au corpus : les lexicalisations qui ne se rencontrent pas ou peu dans le corpus sont éliminées puisqu'elles ne renverront aucun réseau lexical. La collecte des lexicalisations d'un thème est avant tout propre à un corpus. Lors de la projection des lexicalisations sur le corpus, les relations sémantiques en langue qui les structurent sont confrontées à la réalité du discours. Cette comparaison entre langue et discours affine la description des relations sémantiques en mettant au jour des degrés d'intersection sémantique au sein de la relation d'équivalence ou en révélant le brouillage sémantique qui s'opère en discours par rapport à la langue en raison de la déformabilité et de l'idiosyncrasie qui apparaissent en production langagière.

2.2.2 Un corpus homogène et restreint

En raison du caractère exploratoire de la méthode et afin de compenser le grand nombre de résultats susceptibles d'apparaître avec une multitude de lexicalisations, le choix est fait d'un corpus restreint – quelques textes seulement – mais avec une prégnance thématique forte. La prégnance thématique du corpus assure la pertinence des résultats, c'est-à-dire l'apparition, dans le corpus, des lexicalisations choisies. Le nombre de lexicalisations et la prégnance du corpus sont des critères propres à l'étude onomasiologique de la synonymie. Par comparaison, Bourion (1995), qui étudie le thème de la peur, se limite à huit lexicalisations substantives dont la lexicalisation privilégiée *peur*, augmentées des dérivés morphologiques de ces substantifs (adjectifs et verbes). Ce thème restreint, peu fourni en lexicalisations, oblige à recourir à un large corpus : le fonds roman de FRANTEXT pour la période 1830-1970. Ainsi, pour l'étude de la synonymie, les proportions sont inversées : une grande quantité de lexicalisations pour un faible corpus. Vient s'ajouter la prégnance thématique du corpus afin de s'assurer de la pertinence de la recherche menée⁹. Au-delà de la prégnance thématique, le choix du corpus est conditionné par les mêmes impératifs que pour l'analyse thématique : le corpus doit être homogène concernant le type de discours, le genre, et l'état de langue.

3 Le matériau linguistique

3.1 Le corpus de textes

Le corpus, choisi en accord avec les thèmes afin de s'assurer du critère de la prégnance thématique, est constitué de trois textes du XVII^e siècle : les *Caractères* de La Bruyère, les *Maximes et Réflexions diverses* de La Rochefoucauld et les *Pensées* de Pascal. La thématique générale de ces textes est donc la morale. Si La Bruyère et La Rochefoucauld sont qualifiés de moralistes (Lafond, 1992), Pascal est plutôt philosophe (Tourette, 2008). Cela ne l'empêche pas de traiter de thèmes moraux ; d'ailleurs, il est souvent rapproché de La Bruyère et La Rochefoucauld, en raison de l'influence qu'il a exercée sur ces auteurs.

L'homogénéité du corpus repose sur l'état de langue du français classique, puisque ces textes ont tous été publiés durant la seconde moitié du XVII^e siècle. Ils appartiennent au discours constituant littéraire et relèvent du discours argumentatif et du genre essai. Cependant, la forme discontinue de ces trois textes n'en fait pas des essais types. Les textes de La Bruyère et de La Rochefoucauld sont constitués de formes brèves et appartiennent à la catégorie des sentences. Le texte inachevé de Pascal n'appartient pas au genre des sentences, même si sa forme fragmentaire, voire lapidaire, l'en rapproche. Du fait de son inachèvement, le texte de Pascal est constitué de fragments et non de formes brèves (*Ibid.*). Malgré ces différences, les textes sont fédérés par leur discontinuité qui suppose l'autonomie, la concision et la condensation linguistique dans chaque forme brève ou fragment et par la thématique de la morale.

3.2 Les thèmes et les lexicalisations choisies

Dans la thématique générale de la morale, ce sont les thèmes du vice et de la vertu qui ont été choisis. Ce choix s'est fait conjointement en raison de leur indissociabilité ; il n'aurait pas été justifié de choisir un thème sans l'autre, tant ils sont entrelacés, particulièrement dans les écrits des moralistes. Paradoxalement, l'antagonisme qui règne entre ces deux thèmes fait qu'ils sont souvent rapprochés et que leurs lexicalisations peuvent être cooccurrentes. Dans cette étude de la synonymie, l'accent est mis sur la relation d'équivalence, mais les autres relations sémantiques sont aussi à l'honneur : la relation d'hyponymie en raison de la généralité des thèmes choisis, susceptibles de rassembler de nombreux hyponymes, et la relation d'opposition en raison de l'antagonisme qui règne entre les thèmes. Cela conduit à ne pas se focaliser sur la seule relation d'équivalence mais, en prenant en compte les autres relations sémantiques¹⁰, à examiner leurs interactions en discours. Dans la présente étude, seule la relation d'équivalence est traitée.

Comme les textes du corpus datent du XVII^e siècle, la définition des thèmes et des lexicalisations s'est basée sur les dictionnaires de langue du français classique, soit le Furetière (1690) et le premier dictionnaire de l'Académie française (1694). La collecte des lexicalisations a débuté à partir des lexicalisations privilégiées *vice* et *vertu* et s'est étendue à leurs dérivés morphologiques puis à leurs synonymes et leurs dérivés. La récolte des hyponymes est partie des sept péchés capitaux¹¹ pour le vice et des vertus cardinales et théologiques¹² pour la vertu. Mais la lecture des textes du corpus a permis d'infléchir l'accent religieux donné à cette collecte en révélant des lexicalisations de vices et de vertus mondains tels qu'*honnêteté* ou *intérêt*. Enfin ces hyponymes ont eux-mêmes été enrichis par leurs synonymes. La recherche de synonymes s'est effectuée à partir du Dictionnaire Électronique des Synonymes du CRISCO. Comme les relations de synonymie proposées par ce dictionnaire relèvent du français moderne, elles ont été sanctionnées par les dictionnaires de français classique. Enfin, une confrontation des lexicalisations récoltées au corpus a permis d'éliminer les lexicalisations absentes ou trop faiblement fréquentes dans les textes. Ainsi, après la collecte et la validation statistique, subsiste une centaine de lexicalisations de forme substantive auxquelles il faut ajouter les dérivés adjectivaux et adverbiaux ainsi que les flexions des substantifs et des adjectifs. A titre d'exemple, le thème de l'orgueil comporte les lexicalisations *amour-propre*, *arrogance*, *fierté*, *hauteur*, *orgueil*, *présomption*, *suffisance*, *superbe*, *vanité* et leurs dérivés, qui se situent dans une relation d'équivalence forte, l'équivalence sémantique au sens large englobant toutes les autres lexicalisations du thème du vice, telles que *avarice*, *jalousie*, *colère*, *hypocrisie* ou *injustice*.

3.3 Traitement automatique du lexique et des textes

Le corpus et les lexicalisations organisées sous la forme d'une base lexicale ont été structurés au format XML pour rendre possible leur interrogation par une requête en langage XQuery. Les textes sont segmentés en paragraphes, empan textuels dans lesquels s'effectue la recherche des réseaux lexicaux et qui concorde pour l'essentiel avec un découpage en formes brèves. Lors de la première étape de l'expérimentation, seuls les lexicalisations privilégiées et leurs dérivés morphologiques sont projetés sur le corpus ; les paragraphes qui contiennent ces lexicalisations sont récupérés pour la seconde étape de projection qui consiste à projeter l'ensemble du lexique. Enfin sont récupérés les paragraphes qui contiennent une lexicalisation en plus d'une des lexicalisations privilégiées. Ces deux étapes permettent de révéler des réseaux lexicaux comportant au moins deux lexicalisations de la base mais aussi de réduire le bruit dû à la polysémie de certaines unités lexicales. La requête s'appuie sur le module « plein texte » du langage XQuery et peut ainsi convoquer un thesaurus pour effectuer une recherche lexicale de grande échelle sur un corpus. La base lexicale répertoriant les lexicalisations des thèmes étudiés est utilisée afin de détecter les instanciations des thèmes présentes dans les textes du corpus. Cette requête appose un balisage spécifique à toute lexicalisation détectée, ce qui permet de les visualiser et laisse donc la possibilité d'opérer des traitements automatiques sur ces lexicalisations. Les résultats obtenus sont classés en fonction des configurations syntagmatiques prises par les réseaux lexicaux et les relations sémantiques qui les sous-tendent ; le classement se partage en réseaux lexicaux microsyntactiques, c'est-à-dire intra-propositionnels, et en réseaux lexicaux macrosyntaxiques, c'est-à-dire inter-propositionnels (ces réseaux lexicaux pouvant devenir textuels lorsqu'ils dépassent le cadre de la phrase ou de la période).

4 L'équivalence sémantique en parataxe de mots

Parmi les configurations syntagmatiques dont témoignent les réseaux lexicaux, les configurations parataxiques donnent le plus à voir de relations d'équivalence. Au sein de cette configuration, les parataxes de mots parasitent le moins la relation d'équivalence en raison de leur déploiement microsyntactique. De ce fait, les relations sémantiques en discours sont susceptibles d'être proches des relations en langue, tout en laissant place à la déformabilité de la langue, d'où une comparaison possible entre les deux types de relation.

La parataxe est l'un des modes d'organisation de la syntaxe avec l'hypotaxe ; elle concerne les procédés de coordination des propositions ou des constituants des propositions. Contrairement à l'hypotaxe, aucun

rapport de dépendance hiérarchique n'est instauré entre ces constituants ou ces propositions. Ces éléments « restent sur un pied d'égalité syntaxique [...], gardent leur autonomie catégorielle [...] et forment ensemble une unité complexe qui appartient à la même catégorie » (Riegel *et coll.*, 2009 : 871). La parataxe est divisée en deux phénomènes : la coordination et la juxtaposition. « La coordination se distingue de la juxtaposition par la présence d'un ou plusieurs outils de jonction entre les éléments conjoints (les conjonctions de coordination) » (*Ibid.*).

Pour cette étude de l'équivalence sémantique en parataxe de mots, les deux thèmes sont traités séparément et l'attention est portée sur les lexicalisations hyponymiques des thèmes ; les lexicalisations privilégiées sont exclues, afin de conserver le même niveau de spécificité sémantique. Les résultats sont classés en fonction des degrés d'intersection sémantique de la relation d'équivalence puis par thème et par auteur. Les degrés d'intersection sémantique sont au nombre de trois : intersection forte, intermédiaire et faible. L'intersection forte correspond aux cas où les lexicalisations en jeu sont étiquetées comme synonymes dans la base lexicale et sont en relation d'interdéfinition dans les dictionnaires de langue : ces lexicalisations possèdent de nombreux sèmes en commun. L'intersection faible correspond aux cas où les lexicalisations sont cohyponymes dans la base lexicale, ce qui signifie qu'elles n'ont que les sèmes /vice/ ou /vertu/ en commun. L'intersection intermédiaire est une situation médiane : les lexicalisations sont plus que des cohyponymes mais ne sont pas étiquetées comme synonymes dans la base lexicale ; un autre sème que /vice/ ou /vertu/ permet de les relier mais les sèmes qui les distinguent demeurent assez nombreux pour ne pas en faire des lexicalisations entretenant une intersection sémantique forte. Cette intersection est le fait de la créativité linguistique de l'auteur et témoigne de l'idiosyncrasie dans la langue.

4.1 Intersection sémantique forte

4.1.1 Les lexicalisations hyponymiques du vice

La Bruyère

Les résultats de La Bruyère se rapportant à l'intersection sémantique forte sont classés en fonction des sous-thèmes révélés par les parataxes de mots.

Le sous-thème de l'orgueil apparaît dans les combinaisons lexicales suivantes : *fierté* et *hauteur* ou *hauteur* et *fierté*, *vanité* et *présomption* mais aussi entre deux adverbes, *hautement* et *superbement*.

[...] Ici l'on croit se haïr avec plus de *fierté* et de *hauteur*, et peut-être avec plus de dignité [...] (LB, IX, 53)¹³

[...] L'homme, de sa nature, pense *hautement* et *superbement* de lui-même, et ne pense ainsi que de lui-même [...] (LB, XI, 69)

[...] vice honteux, et qui par son excès rentre toujours dans la *vanité* et dans la *présomption*, et ne persuade pas tant à celui qui en est blessé qu'il a plus d'esprit et de mérite que les autres, qu'il lui fait croire qu'il a lui seul de l'esprit et du mérite. (LB, XI, 85)

[...] quelle misère ! et puisqu'il est vrai que dans un si étrange commerce, ce que l'on pense gagner d'un côté on le perd de l'autre, ne reviendrait-il pas au même de renoncer à toute *hauteur* et à toute *fierté*, qui convient si peu aux faibles hommes, et de composer ensemble, de se traiter tous avec une mutuelle bonté, qui, avec l'avantage de n'être jamais mortifiés, nous procurerait un aussi grand bien que celui de ne mortifier personne ? (LB, XI, 131)

Le rapprochement parataxique suppose une grande affinité de sens entre les lexicalisations, au point qu'elles sont le plus souvent en relation d'interdéfinition, d'après les dictionnaires du français classique. Cependant, la coordination laisse supposer que des différences de sens distinguent ces mots. Il aurait été

redondant de coordonner des mots de même sens, au-delà de la rupture avec le principe d'économie que cela présume. Mais, dans ces coordinations, les différences ne sont pas explicitées, et c'est l'intersection sémantique forte qui ressort de prime abord.

La *paresse*, l'indolence et l'*oisiveté*, vices si naturels aux enfants, disparaissent dans leurs jeux, où ils sont vifs, appliqués, exacts, amoureux des règles et de la symétrie, où ils ne se pardonnent nulle faute les uns aux autres, et recommencent eux-mêmes plusieurs fois une seule chose qu'ils ont manquée [...] (LB, XI, 55)

Pour le sous-thème de la paresse, la parataxe associe les lexicalisations *paresse* et *oisiveté* complétées par le corrélat *indolence*, qui pourra être ajouté à la base lexicale. Ce cas témoigne de l'utilité des corrélatés pour enrichir la base lexicale¹⁴.

Je ne sais s'il y a rien au monde qui coûte davantage à approuver et à louer que ce qui est plus digne d'approbation et de louange, et si la vertu, le mérite, la beauté, les bonnes actions, les beaux ouvrages, ont un effet plus naturel et plus sûr que l'*envie*, la *jalousie*, et l'antipathie. [...] (LB, XII, 8)

Pour le sous-thème de la jalousie, les lexicalisations principales, *envie* et *jalousie*, sont coordonnées et complétées par un corrélat, *antipathie* qui ne possède pas le sème /convoitise/ comme *envie* et *jalousie*. Mais, puisque ce corrélat peut être coordonné avec les lexicalisations de la jalousie, cela suppose que le sentiment haineux intrinsèque à l'antipathie fait partie de la jalousie. Ainsi, le sous-thème moral de la jalousie a partie liée avec le sous-thème émotionnel de la haine, ou du moins on peut considérer que ces sous-thèmes sont voisins.

C'est une pratique ancienne dans les cours de donner des pensions et de distribuer des grâces à un musicien, à un maître de danse, à un farceur, à un joueur de flûte, à un *flatteur*, à un *complaisant*: ils ont un mérite fixe et des talents sûrs et connus qui amusent les grands et qui les délassent de leur grandeur [...] (LB, XIII, 28)

Les dérivés du sous-thème de la flatterie, *flatteur* et *complaisant*, apparaissent juxtaposés à la fin d'une énumération. Les dictionnaires du français classique tendent à l'interdéfinition, puisque *flatteur* apparaît dans la définition de *complaisant* et inversement. Toutefois, une nuance transparait entre ces deux lexicalisations : le sème /vice/ est plus fort dans *flatterie* que dans *complaisance* car *complaisance* peut parfois être connoté positivement, ce qui n'est jamais le cas de *flatterie*.

[...] tout le *bas*, tout le *faible* et tout l'indigne s'y trouvent. [...] (LB, IX, 53)

Le dernier sous-thème rencontré parmi les lexicalisations hyponymiques du vice est celui de la bassesse manifesté par les dérivés *bas* et *faible*. Vient s'y ajouter le corrélat *indigne* qui appuie le sémantisme de *bas* et oriente la définition du sous-thème vers la bassesse et non la faiblesse.

4.1.2 Les lexicalisations hyponymiques de la vertu

La Bruyère

On retrouve à quatre reprises chez La Bruyère la coordination de *modération* et *sagesse* que ce soit sous leur forme lemmatique ou en tant que dérivés morphologiques.

[...] ils les imitent, ils se tiennent fort au-dessus de la gravité de la robe, et se croient dispensés par leur âge et par leur fortune d'être *sages* et *modérés*. [...] (LB, VII, 7)

[...] Quelques pas que quelques-uns fassent par vertu vers la *modération* et la *sagesse*, un premier mobile d'ambition les emmène avec les plus avarés, les plus violents dans leurs désirs et les plus ambitieux [...] (LB, VIII, 22)

L'esprit de *modération* et une certaine *sagesse* dans la conduite laissent les hommes dans l'obscurité : il leur faut de grandes vertus pour être connus et admirés, ou peut-être de grands vices. (LB, XII, 112)

[...] quelle force invincible et accablante des témoignages rendus successivement et pendant trois siècles entiers par des millions de personnes les plus *sages*, les plus *modérées* qui fussent alors sur la terre, et que le sentiment d'une même vérité soutient dans l'exil, dans les fers, contre la vue de la mort et du dernier supplice ! [...] (LB, XVI, 34)

La modération intervient dans la définition de la sagesse à côté d'autres vertus comme la prudence. L'interdéfinition vient appuyer l'interprétation de cette coordination en faveur de l'intersection sémantique forte.

[...] un esprit de *droiture* et d'*équité* qui fait qu'on le suit jusques à prononcer quelquefois contre soi-même en faveur du peuple, des alliés, des ennemis ; [...] (LB, X, 35)

Les lexicalisations *droiture* et *équité* sont coordonnées pour former le complément du nom *esprit*. Cette coordination repose sur une forte intersection sémantique entre ces lexicalisations, puisqu'elles sont en relation d'interdéfinition : dans les dictionnaires de français classique, la lexicalisation *équité* est utilisée pour définir *droiture*.

Pascal

[...] Voilà les sentiments qui naîtraient d'un cœur qui serait plein d'*équité* et de *justice*.
[...] (P, 758)

Cet extrait donne à voir une coordination de deux lexicalisations du thème de la vertu : *équité* et *justice*. Ces lexicalisations appartiennent au sous-thème de la justice, si bien qu'elles entretiennent une relation d'équivalence approchée basée sur une intersection sémantique forte. D'ailleurs, *équité* est définie à l'aide de *justice* par les dictionnaires de langue du français classique. Ces lexicalisations ont essentiellement le sème /justice/ en commun, ce qui fait de l'équité un type de justice particulier : selon Furetière, elle possède en plus le sème /adoucissement/.

Les intersections sémantiques fortes donnent à voir des parataxes entre des mots de sens très proche, voire en relation d'interdéfinition étant donné leurs sèmes communs¹⁵. De ce fait, cela laisse peu de place à la recherche de différences sémantiques. Les lexicalisations présentes dans ces intersections correspondent à ce qui est étiqueté comme synonyme dans la base lexicale et ce qui est communément répertoriés comme tel dans les dictionnaires. Ainsi, le discours est une reproduction des relations répertoriées en langue, il n'y a donc pas déphasage entre la langue et le discours avec ce type d'intersection.

4.2 Intersection sémantique intermédiaire

4.2.1 Les lexicalisations hyponymiques du vice

La Bruyère

La bassesse, la flatterie et la fourberie sont trois vices contractés par les courtisans qui cherchent à parvenir, ils sont tous trois avilissants pour eux.

[...] qui bannirait des cours les brigues, les cabales, les mauvais offices, la *bassesse*, la *flatterie*, la *fourberie* [...] (LB, II, 11)

Les noms de ces vices ferment une énumération commencée par trois noms d'actions indiquant les manœuvres effectuées par les courtisans pour parvenir. Ainsi, ces vices semblent faire pendant à ces trois actions. En avilissant le courtisan, ils le poussent à manœuvrer contre autrui pour qu'il arrive à ses fins, d'où le lien étroit entre ces vices et les actions mentionnées. Les manœuvres et l'avilissement du courtisan prêt à tout pour son avancement fédèrent ces lexicalisations et constituent l'intersection sémantique. Il s'agit bien d'une intersection intermédiaire : une intersection faible se serait contentée du sème /vice/ comme sème commun et une intersection forte aurait nécessité davantage de sèmes communs. Si l'avilissement et les manœuvres rapprochent ces trois lexicalisations, trop de différences sémantiques les

distinguent encore : la bassesse fait que l'on s'incline devant autrui, la flatterie permet de s'attirer les bonnes grâces d'autrui pour être bien vu et la fourberie conduit à tromper autrui, à dissimuler ses intentions.

[...] ne savent-ils pas justifier les mauvais succès par les bonnes intentions, prouver la bonté d'un dessein et la justesse des mesures par le bonheur des événements, s'élever contre la *malignité* et l'*envie* pour accorder à de bonnes entreprises de meilleurs motifs, donner des explications favorables à des apparences qui étaient mauvaises, détourner les petits défauts, ne montrer que les vertus, et les mettre dans leur jour, semer en mille occasions des faits et des détails qui soient avantageux, et tourner le ris et la moquerie contre ceux qui oseraient en douter ou avancer des faits contraires ? [...] (LB, IX, 34)

L'intersection sémantique entre *malignité* et *envie* est de type intermédiaire. Elle repose sur l'idée de nuisance commune à ces vices : le malin nuit à son prochain et l'envieux cherche à posséder le bien de son prochain au point de lui nuire pour l'obtenir. Même si le sème /nuisance/ n'a pas la même valeur pour l'un et l'autre vice – essentielle pour la *malignité*, télique pour l'*envie* – ce sème est bien commun aux deux vices et vient s'ajouter au sème /vice/, ce qui en fait une intersection intermédiaire.

[...] Adraste était si *corrompu* et si *libertin*, qu'il lui a été moins difficile de suivre la mode et se faire dévot : il lui eût coûté davantage d'être homme de bien. (LB, XI, 147)

Les dérivés *corrompu* et *libertin* ne sont pas en relation d'interdéfinition. *Libertin* prend une connotation religieuse car, dans les dictionnaires du français classique, le *libertin* est celui qui n'obéit pas aux lois de l'Église et qui en fait profession. Cette connotation prend d'autant plus de force que *libertin* est rapproché de *dévo*t, ce qui déprécie d'autant plus *dévo*t. L'intersection sémantique est donc de type intermédiaire et elle repose sur le sème /débauche/ en plus du sème /vice/ commun à ces deux lexicalisations.

Quand un courtisan sera humble, guéri du faste et de l'ambition ; qu'il n'établira point sa fortune sur la ruine de ses concurrents ; qu'il sera équitable, soulagera ses vassaux, payera ses créanciers ; qu'il ne sera ni *fourbe* ni *médisant* [...] (LB, XIII, 23)

Les adjectifs *fourbe* et *médisant* coordonnés par *ni* manifestent les caractéristiques du courtisan. On retrouve la fourberie, mais cette fois, la médianse vient s'y ajouter. Comme la flatterie déjà coordonnée à la fourberie, la médianse se trouve éloignée de la vérité : si la flatterie consiste à dire plus de bien qu'il ne faut d'une personne, la médianse consiste à en dire plus de mal qu'il ne faut. Ainsi, la fourberie et la médianse sont deux vices apparentés par leur éloignement de la vérité : la fourberie pousse à tromper, c'est-à-dire à agir avec fausseté et la médianse amène à parler fausement. C'est donc la fausseté du courtisan qui est en jeu ici, et elle transparait plus précisément par ses paroles et ses actions. Il est bien question d'une intersection intermédiaire, puisque le sème /fausseté/ s'ajoute au sème /vice/, mais ces deux vices s'appliquent à des sphères différentes : le dire et l'agir.

La Rochefoucauld

[...] il est incapable d'*envie* ni d'*avarice*, soit par vertu ou par inapplication. [...] (LR, portrait du Cardinal de Retz)

Dans le portrait du cardinal de Retz par La Rochefoucauld, les lexicalisations *envie* et *avarice* sont coordonnées. Il s'agit d'une intersection intermédiaire car ces deux vices ont en commun le sème /possession/ en plus du sème /vice/ : l'avare est celui qui conserve farouchement son bien et l'envieux est celui qui jalouse le bien d'autrui.

Pascal

Car ne voyant pas la vérité entière, ils n'ont pu arriver à une parfaite vertu ; les uns considérant la nature comme incorrompue, les autres comme irréparable, ils n'ont pu fuir ou l'*orgueil* ou la *paresse* qui sont les deux sources de tous les vices, puisqu'il ne peut sinon ou s'y abandonner par lâcheté, ou en sortir par l'*orgueil*. (P, 194)

[...] Comme les deux sources de nos péchés sont l'*orgueil* et la *paresse*, Dieu nous a découvert deux qualités en lui pour les guérir, sa *miséricorde* et sa *justice*. [...] (P, 648)

Dans ces deux extraits, Pascal rapproche les noms de deux vices, *orgueil* et *paresse*, dans une coordination alternative pour le premier cas et additive pour le second car il en fait « les deux sources de tous les vices » ou « les deux sources de nos péchés ». Ces lexicalisations ont en commun le sème /vice/ mais également le sème /source/, ce qui fait de leur rapprochement une intersection intermédiaire. Le caractère premier de ces vices confirme leur statut de péchés capitaux, péchés suprêmes qui entraînent les autres types de péchés.

4.2.2 Les lexicalisations hyponymiques de la vertu

La Bruyère

La Bruyère coordonne les lexicalisations *bravoure* et *libéralité* et reproduit cette parataxe à la fin de l'extrait mais en recourant aux dérivés morphologiques *brave* et *libéral*.

Il est vrai qu'il y a deux vertus que les hommes admirent, la *bravoure* et la *libéralité*, parce qu'il y a deux choses qu'ils estiment beaucoup, et que ces vertus font négliger, la vie et l'argent : aussi personne n'avance de soi qu'il est *brave* ou *libéral*. (LB, XI, 84)

Ces deux vertus conduisent à négliger des choses que l'on estime, en l'occurrence l'argent et la vie, ce qui peut être vu comme une prodigalité de ses biens ou de sa vie. Le sème /prodigalité/ renforce le rapprochement entre ces deux vertus en venant s'ajouter au sème /vertu/ déjà commun aux deux lexicalisations. Il est question d'une intersection intermédiaire, puisque la *bravoure* et la *libéralité* s'appliquent à deux sphères distinctes, les biens et la vie.

La Rochefoucauld

Ce que nous prenons pour des vertus n'est souvent qu'un assemblage de diverses actions et de divers intérêts, que la fortune ou notre industrie savent arranger ; et ce n'est pas toujours par *valeur* et par *chasteté* que les hommes sont *vaillants*, et que les femmes sont *chastes*. (LR, 1)

La Rochefoucauld recourt à deux lexicalisations distinctes pour désigner la vertu principale de tout un chacun : la *valeur* pour les hommes et la *chasteté* pour les femmes. Le rapprochement repose sur les sèmes /vertu/ et /honneur/ communs aux deux lexicalisations. Furetière définit l'un des sens d'*honneur* comme suit : « S'applique plus particulièrement à deux sortes de vertus : à la vaillance pour les hommes, et à la chasteté pour les femmes ». La Rochefoucauld pointe en réalité l'honneur au travers de ces deux vertus mais il est obligé de recourir à deux lexicalisations distinctes car la vertu d'honneur change de nom en fonction du genre des individus. Ainsi, le réseau lexical qui s'élabore à partir de cette coordination repose sur une intersection sémantique intermédiaire.

Pascal

[...] Comme les deux sources de nos péchés sont l'*orgueil* et la *paresse*, Dieu nous a découvert deux qualités en lui pour les guérir, sa *miséricorde* et sa *justice*. [...] (P, 648)

Les lexicalisations *miséricorde* et *justice* en parataxe de mots entretiennent une relation d'équivalence sémantique par intersection intermédiaire. Si la *justice* est le fait de rendre à chacun ce qui lui appartient, la *miséricorde* intervient à la fin du processus de *justice* puisqu'elle est pardon, grâce accordée à celui qui a fauté. Elles s'insèrent toutes deux dans un processus juridique et ce sème commun facilite leur rapprochement, d'où une intersection sémantique intermédiaire.

Les lexicalisations des intersections intermédiaires constituent un stade intermédiaire entre ce qui est appelé *synonyme* et les cohyponymes dans la mesure où elles possèdent un sème commun en plus du sème manifestant le genre commun, mais leurs différences sémantiques sont trop nombreuses pour

qu'elles soient en relation d'interdéfinition comme dans les intersections fortes. Elles sont le plus souvent rapprochées en contexte par l'auteur, ce rapprochement reposant sur une afférence contextuelle ou socialement normée. Ainsi elles trouvent leur justification dans le contexte discursif en raison d'associations sémantiques propres au locuteur. Elles témoignent de la différence qui réside entre langue et discours et entre les relations sémantiques répertoriées dans l'un ou l'autre plan du langage. Elles manifestent donc l'espace de créativité linguistique que le discours octroie au locuteur par rapport aux normes constituées par la langue. L'analyse thématique, de par sa large couverture lexicale et donc de par sa souplesse, rend possible la mise au jour des rapprochements lexicaux propres aux auteurs, et ainsi met en relation sémantique lexicale et stylistique¹⁶. Elle montre aussi qu'il est nécessaire de ne pas seulement s'en tenir à la langue, qui s'avère plus restrictive, dès lors qu'on étudie les relations sémantiques ; de plus, elle montre que la relation d'équivalence sémantique ne se limite pas à la seule synonymie.

4.3 Intersection sémantique faible

4.3.1 Les lexicalisations hyponymiques du vice

La Bruyère

[...] et elles se perdent tristement par la *présomption* et par l'*envie*. (LB, III, 43)

La Bruyère épingle les femmes coquettes qui se sont fait dévotes pour suivre la mode. Ces deux attitudes opposées sont impulsées par des vices distincts : la coquetterie est le fait de la galanterie, de la bonne chère et l'oisiveté, alors que la dévotion est le fait de la présomption et de l'envie. Ces deux vices sont rapprochés car ils s'appliquent aux femmes dévotes et expliquent leur comportement. Mais il n'est pas possible de déceler un sème commun à ces lexicalisations autre que le sème /vice/ ; ainsi, s'agit-il d'une intersection faible due à un rapprochement dépendant du contexte.

[...] Voulez-vous qu'ils empiètent sur celui des gens de bien, qui avec les vices cachés fuient encore l'*orgueil* et l'*injustice* ? (LB, XII, 22)

Les lexicalisations *orgueil* et *injustice* sont rapprochées, mais aucun sème commun, en dehors de /vice/, n'est décelable. Le contexte permet de rapprocher ces vices : ils sont également fuis par les gens de bien, tout comme les vices cachés. Il s'agit donc d'une intersection faible.

[...] Quelques pas que quelques-uns fassent par vertu vers la modération et la sagesse, un premier mobile d'ambition les emmène avec les plus *avares*, les plus violents dans leurs désirs et les plus *ambitieux* [...] (LB, VIII, 22)

La parataxe de mots relie les dérivés *avares* et *ambitieux* ainsi que le corrélat *violents dans leurs désirs*. On pourrait déceler /argent/ comme sème commun à ces lexicalisations en plus de /vice/ ce qui donnerait lieu à une intersection intermédiaire. Cependant, le corrélat *violents dans leurs désirs* placé entre les deux lexicalisations annule cette interprétation. Ainsi, si aucun sème commun ne vient s'ajouter à /vice/ dans ce rapprochement, cette parataxe témoigne seulement d'une intersection faible. C'est le « premier mobile d'ambition » opposé à la modération et à la sagesse qui permet ce rapprochement de noms d'agents vicieux.

En raison du nombre de lexicalisations et de corrélats qu'elles alignent, les énumérations favorisent l'intersection faible : plus il y a de mots en parataxe, plus il devient difficile de saisir un sème commun en plus du sème général /vice/.

[...] ils prennent de la cour ce qu'elle a de pire : ils s'approprient la *vanité*, la mollesse, l'*intempérance*, le *libertinage*, comme si tous ces vices leur étaient dus, et, affectant ainsi un caractère éloigné de celui qu'ils ont à soutenir, ils deviennent enfin, selon leurs souhaits, des copies fidèles de très méchants originaux. (LB, VII, 7)

L'incivilité n'est pas un vice de l'âme, elle est l'effet de plusieurs vices: de la sottise *vanité*, de l'ignorance de ses devoirs, de la *paresse*, de la stupidité, de la distraction, du mépris des autres, de la *jalousie*. [...] (LB, XI, 8)

Dans le premier de ces deux extraits, l'énumération contient *vanité*, *intempérance* et *libertinage* auxquels on peut ajouter le corrélat *mollesse*. Cette énumération pointe les pires vices de la Cour : on peut donc déceler le sème /intensité/ mais ce sème ne constitue pas un point commun supplémentaire entre ces vices, il est inhérent au sémème « vice ». On a donc affaire à une intersection faible plutôt qu'à une intersection intermédiaire.

Dans le second cas, sont énoncés les vices qui constituent l'incivilité, d'où une énumération. Apparaissent les lexicalisations *vanité*, *paresse* et *jalousie*, ainsi que de nombreux corrélat. Les lexicalisations et les corrélat renvoient à de trop nombreux sous-thèmes pour déceler un sème commun à tous en dehors de /vice/, d'où une intersection sémantique faible.

La Rochefoucauld

Cette clémence dont on fait une vertu se pratique tantôt par *vanité*, quelquefois par *paresse*, souvent par crainte, et presque toujours par tous les trois ensemble. (LR, 16)

La lexicalisation *vanité* est coordonnée à la lexicalisation *paresse*. Seul le sème /vice/ rapproche les deux lexicalisations, d'où une intersection sémantique faible. La justification de cette parataxe se trouve dans le contexte : la vanité, la paresse et la crainte sont les trois ingrédients de la clémence chez La Rochefoucauld. Sa définition pessimiste s'appuie sur des vices ou un sentiment négatif, voire les trois en même temps.

Pascal

[...] Si l'on ne se connaît plein de *superbe*, d'*ambition*, de *concupiscence*, de *faiblesse*, de misère et d'*injustice*, on est bien aveugle. [...] (P, 507)

Cet extrait donne à voir une énumération de vices coordonnés par la conjonction *et* : *superbe*, *ambition*, *concupiscence*, *faiblesse* et *injustice*, auxquels vient s'ajouter le corrélat *misère*. Ces lexicalisations sont réunies par une relation d'équivalence reposant sur une intersection sémantique faible. Pascal énumère différents vices, sans que l'on puisse déterminer le sous-thème précis qui les rassemble. Aussi, ont-ils en commun le seul sème /vice/. Ce réseau lexical repose sur une intersection sémantique faible se rapportant à une relation de cohyponymie, puisqu'il énonce les principaux vices dont souffrent les hommes.

[...] Ainsi le juste ne prend rien pour soi du monde, ni des applaudissements du monde, mais seulement pour ses passions, desquelles il se sert comme maître en disant à l'une : « Va et viens. » « Sub te erit appetitus tuus. » Ses passions ainsi dominées sont vertus ; l'*avarice*, la *jalousie*, la *colère*, Dieu même se les attribue. Et ce sont aussi bien vertus que la clémence, la pitié, la constance, qui sont aussi des passions. (P, 517)

Pascal énumère des vices par juxtaposition : *avarice*, *jalousie* et *colère*. Ce rapprochement procède d'une intersection sémantique faible puisque ces trois lexicalisations n'ont en commun que le sème /vice/. Toutefois, on pourrait y voir une intersection sémantique intermédiaire, puisque le sème /haine/ est commun *jalousie* et *colère* ou le sème /incontrôlable/ commun aux trois. Mais justement, le propre des vices est de ne pouvoir être contrôlés. C'est d'ailleurs le propos du fragment : pour Pascal, des vices qui parviennent à être contrôlés peuvent devenir des vertus. Le rapprochement repose sur une intersection sémantique faible.

4.3.2 Les lexicalisations hyponymiques de la vertu

La Bruyère

[...] Ce Tryphon qui a tous les vices, je l'ai cru *sobre*, *chaste*, *libéral*, *humble* et même *dévo*t [...] (LB, VI, 50)

[...] Mais si vous êtes *sage, tempérant, modeste*, civil, généreux, reconnaissant, laborieux, d'un rang d'ailleurs et d'une naissance à donner des exemples plutôt qu'à les prendre d'autrui, et à faire les règles plutôt qu'à les recevoir, convenez avec cette sorte de gens de suivre par complaisance leurs dérèglements, leurs vices et leur folie, quand ils auront, par la déférence qu'ils vous doivent, exercé toutes les vertus que vous chérissez [...] (LB, IX, 2)

[...] imiter sa *bonté*, sa docilité, son *équité*, sa vigilance, son intrépidité ? [...] (LB, X, 24)

Rien n'engage tant un esprit raisonnable à supporter tranquillement des parents et des amis les tords qu'ils ont à son égard, que la réflexion qu'il fait sur les vices de l'humanité, et combien il est pénible aux hommes d'être *constants, généreux*, fidèles, d'être touchés d'une amitié plus forte que leur intérêt. [...] (LB, XI, 28)

[...] Celui qui dit froidement de soi, et sans croire blesser la modestie, qu'il est bon, qu'il est *constant*, fidèle, *sincère, équitable*, reconnaissant, n'ose dire qu'il est vif, qu'il a les dents belles et la peau douce : cela est trop fort. (LB, XI, 84)

La jalousie et l'émulation s'exercent sur le même objet, qui est le bien ou le mérite des autres : avec cette différence, que celle-ci est un sentiment volontaire, *courageux, sincère*, qui rend l'âme féconde, qui la fait profiter des grands exemples, et la porte souvent au-dessus de ce qu'elle admire [...] (LB, XI, 85)

L'ampleur de la parataxe est inversement proportionnelle au nombre de sèmes communs aux éléments constituants cette parataxe. Avec les énumérations, il y a plus de chances d'avoir affaire à des intersections faibles. Ici, l'intersection faible repose sur le sème /vertu/. On ne rencontre pas moins de six énumérations de vertus en intersection faible, dont certaines incluent des corrélats. On recense les lexicalisations et les dérivés adjectivaux suivants : *bonté, chaste, constant, courageux, dévot, équité, équitable, généreux, humble, libéral, modeste, sage, sincère, sobre* et *tempérant*. Ces énumérations d'adjectifs désignant des vertus prennent un caractère hyperbolique ; par son exagération, l'hyperbole conduit à l'ironie : La Bruyère se moque de personnages vicieux dont on aurait cru qu'ils étaient vertueux.

Beaucoup de parataxes des lexicalisations en intersection faible apparaissent dans des énumérations, ce qui montre une certaine affinité entre ce type d'intersection sémantique et cette configuration parataxique, affinité qui peut s'expliquer par le nombre de lexicalisations présentes dans une énumération. Si les intersections sémantiques faibles sont contextuelles car certaines lexicalisations sont rapprochées par rapport à d'autres, elles reposent malgré tout sur une relation en langue : la cohyponymie, et ne donnent à voir qu'un seul sème générique commun, le sème /vice/ ou le sème /vertu/ ; leur rapprochement s'effectue par l'intermédiaire d'un terme tiers, en l'occurrence, leur hyperonyme commun. Cela rejoint la définition de la synonymie chez Aristote, synonymie qui est en fait une cohyponymie dans la terminologie moderne. Si la cohyponymie ne jouit pas de la même renommée que la synonymie, elle constitue malgré tout une relation sémantique répertoriée en langue dans des ontologies ou des thesaurus. Ainsi, comme l'intersection forte, l'intersection faible en discours constitue une reproduction de la langue, autrement dit, il n'y a pas déphasage entre les deux plans du langage. Mais cette intersection a l'avantage de montrer l'importance de la cohyponymie en discours, relation sémantique souvent éclipsée par la synonymie. Elle témoigne donc de l'intérêt des ontologies et de thesaurus face aux dictionnaires de synonymes pour relever et étudier le déploiement des thèmes dans un texte.

5 Conclusion

Les réseaux lexicaux adossés à des intersections fortes ou faibles confirment les relations sémantiques validées en langue, même si le discours intervient pour spécifier les cohyponymes ou les synonymes qui peuvent être rapprochés en contexte. En revanche, l'intersection intermédiaire dépend exclusivement du contexte et du locuteur ; elle n'est pas prévisible en langue puisqu'elle repose sur des sèmes afférents.

Elle constitue une partie de l'espace de créativité sémantique que se réserve le locuteur dans sa production langagière, même s'il est limité par les normes sociales et le contexte discursif. Cette intersection sémantique montre qu'il est possible de créer des relations sémantiques en dehors de celles qui sont répertoriées en langue, elle est donc nécessairement ancrée dans le discours et montre l'immanence du sens au texte. La créativité sémantique témoigne de la déformabilité et de la subjectivité dans la langue. Elle se manifeste également dans les rapprochements sémantiques qui brouillent les frontières des relations sémantiques établies en langue, comme faire de deux antonymes des synonymes. En effet, dans le corpus étudié, les relations d'équivalence apparaissent dans des configurations hypotaxes et contribuent à rendre synonymes ou au moins à effectuer des rapprochements entre les lexicalisations antagonistes *vice* et *vertu* (Doualan, 2015). La déformabilité ne touche pas seulement la relation d'équivalence et les réseaux lexicaux constituent une entrée pertinente dans les textes pour faire apparaître les phénomènes de déformabilité ou l'idiosyncrasie dans la langue.

La méthode présentée ici sert en premier lieu à révéler la présence d'un thème donné dans un texte. Appliquée à un corpus, elle permet de sélectionner les textes pertinents par rapport à ce thème et peut donc avoir diverses finalités en didactique pour la recherche automatique de textes portant sur un sujet ou un vocabulaire donné qui sera étudié en classe ; elle peut également servir en ingénierie documentaire pour la recherche ou l'indexation de documents dans des bases de données. La méthode proposée s'avère plus riche qu'une recherche par mots-clés puisqu'elle convoque les unités lexicales sémantiquement proches, ce qui élargit le champ d'investigation mais atténue également les effets de la polysémie des mots-clés puisqu'ils sont recherchés avec des mots qui leur sont sémantiquement liés dans un thème donné.

Références bibliographiques

Corpus de travail

LA BRUYERE J. (de) (1696). *Les Caractères ou les mœurs de ce siècle*. Paris : E. Michallet.

LA ROCHEFOUCAULD F. (de) (1665). *Réflexions ou sentences et maximes morales*. Paris : Barbin.

PASCAL B. (1998). *Œuvres complètes*. M. Le Guern (éd.). Paris : Gallimard.

Références secondaires

ACADEMIE FRANCAISE (1694). *Dictionnaire de l'Académie française*. Paris : Vve J.-B. Coignard & J.-B. Coignard.

ARISTOTE (2002), Catégories, I, Présentation, traduction et commentaires de F. Ildefonse et J. Lallot, Paris, Seuil.

BALDINGER K. (1984). *Vers une sémantique moderne*. Paris : Klincksieck.

BOURION E. (1995). Le réseau associatif de la peur. *L'analyse thématique des données textuelles. L'exemple des sentiments*. F. Rastier (dir.), Paris : Didier Erudition, 107-145.

BURIDANT C. (1980). Les binômes synonymiques : esquisse d'une théorie des couples de synonymes du Moyen-Age au XVIIe siècle. *Bulletin du Centre d'Analyse du Discours*, n°4. Lille : Presses Universitaires de Lille, 5-79.

CULIOLI A. (1990). *Pour une linguistique de l'énonciation. Opérations et représentations*. Tome 1. Paris : Ophrys.

DOULAN G. (2015). *Etude historique, épistémologique et descriptive de la synonymie*. Thèse de doctorat. Université Paris-Sorbonne.

DUCHÁČEK O. (1967). *Précis de sémantique française*. Brno : Universita J. E. Purkyně.

ERLICH D. (1995). Une méthode d'analyse thématique. Exemples de l'ennui et de l'ambition. *L'analyse thématique des données textuelles. L'exemple des sentiments*, F. Rastier (dir.), Paris : Didier Erudition, 85-103.

- FURETIERE A. (1690). *Dictionnaire universel, contenant generalement tous les mots françois tant vieux que modernes, et tous les termes de toutes les Sciences et des Arts*, La Haye/Rotterdam : A. & R. Leers.
- GIRARD G. (1718). *La Justesse de la langue françoise ou Les différentes significations des mots qui passent pour synonymes*. Paris : Veuve d'Houry.
- HONESTE M.-L. (2007). Entre ressemblance et différence : synonymie et cognition, *Le Français moderne*, « *La synonymie en question, Echanges entre les époques* » Tome LXXV, n°1, 160-174.
- LAFAYE P.-B. (1858). *Dictionnaire des synonymes de la langue française*. Paris : Hachette.
- LAFOND J. (dir.) (1992). *Moralistes du XVII^e siècle*. Paris : Laffont.
- LYONS J. (1970). *Linguistique générale : introduction à la linguistique théorique*. Paris : Larousse, trad. F. Dubois-Charlier & D. Robinson.
- MARTIN E. (1995). Thème d'étude, étude de thème. *L'analyse thématique des données textuelles. L'exemple des sentiments*, F. Rastier (dir.), Paris : Didier Erudition, 13-24.
- RASTIER F. (1987). *Sémantique interprétative*. Paris : PUF.
- RASTIER F., CAVAZZA M. & ABEILLE A. (1994). *Sémantique pour l'analyse. De la linguistique à l'informatique*. Paris : Masson.
- RASTIER F. (dir.) (1995). *L'analyse thématique des données textuelles. L'exemple des sentiments*. Paris : Didier Erudition.
- RASTIER F. (2001). *Arts et sciences du texte*. Paris : PUF.
- RIEGEL M., RIOUL R. & PELLAT J.-C. (2009). *Grammaire méthodique du français*. Paris : PUF.
- SOUTET O. (2011). *La Linguistique*. Paris : PUF.
- TOURRETTE É. (2008). *Les formes brèves de la description morale. Quatrains, maximes, remarques*. Paris : Champion.
- ULLMANN S. (1967). *The principles of semantics*. Oxford: Blackwell.

¹ On doit la première définition de la synonymie à Aristote (2002) ; cette synonymie pourrait être assimilée à ce que l'on nomme *cohyponymie* aujourd'hui, puisqu'elle est une relation entre deux espèces d'un même genre : il y a donc synonymie entre *homme* et *bœuf*. Seul l'énoncé de l'essence qui est commun rend possible le rapprochement ; aucune recherche de différences sémantiques n'est opérée. A son origine, la synonymie se situe du côté de l'onomasiologie et non de la sémasiologie.

² La tradition synonymiste, initiée par Girard (1718) au XVIII^e siècle, s'achève au XIX^e siècle avec Lafaye (1858).

³ Parmi les structuralistes qui se sont intéressés à la synonymie et qui l'ont définie de manière différentialiste, on peut nommer Ullmann (1967), Lyons (1970), Ducháček (1967) & Baldinger (1984).

⁴ La synonymie est peu traitée en sémantique cognitive, au détriment d'autres notions comme la polysémie ; on trouve toutefois les travaux d'Honeste (2007) qui procèdent à des substitutions de synonymes en contexte, ce qui revient à rechercher des différences sémantiques.

⁵ Cependant, si l'analyse se concentre sur les équivalences, les différences sémantiques ne sont pas niées, mais seulement écartées de cette analyse.

⁶ Le processus rédactionnel repose essentiellement sur les opérations de contextualisation (ancrage énonciatif et déictique du texte), de planification (organisation des connaissances et des informations) et de textualisation (contrôle des faits de modalisation, de cohésion et de segmentation du texte).

⁷ Les relations sémantiques sont établies entre des unités lexicales de même catégorie grammaticale, or, dans les textes, les réseaux lexicaux s'appuient sur des instanciations lexicales de toutes catégories. La catégorie grammaticale est l'un des critères de choix et même de définition de la synonymie tant que cette notion est étudiée à l'aide du test de la substitution en contexte, puisqu'elle en est un prérequis.

⁸ La démarche onomasiologique du lexicologue consiste à « délimiter un champ notionnel auquel on fait correspondre un ensemble structuré de mots, constituant un champ lexical » (Soutet, 2011 : 264). Mais « pour que

l'étude d'un champ soit possible, c'est-à-dire l'examen de la hiérarchisation et des corrélations de ses constituants, encore faut-il que la liste de ceux-ci soit clairement et exhaustivement délimitée » (*Ibid.*). Si l'exhaustivité n'est pas possible, au moins faut-il tendre vers elle. Il faut donc avoir conscience du flou qui entoure un thème et sa couverture lexicale lorsque l'on entreprend une analyse thématique.

⁹ Une fois l'intérêt de cette méthode exploratoire révélé – une mise au jour efficiente d'un thème dans un corpus –, elle pourra être appliquée à d'autres corpus sans qu'il y ait restriction de la taille du corpus ou qu'il y ait prégnance thématique. D'ailleurs, cette méthode aura comme objectif possible la sélection de textes ou d'extraits de textes manifestant un thème au sein d'un corpus.

¹⁰ Une étude comparée des relations sémantiques montrent l'inconsistance de la synonymie qui se fonde seulement sur la communauté de sens. Si les autres relations sémantiques – hyperonymie et antonymie essentiellement – reposent sur un fond de sens commun, elles se singularisent par les différences de sens qui leur sont propres : l'hyponyme possède un trait sémantique supplémentaire par rapport à son hyperonyme et les antonymes se distinguent par leurs traits sémantiques antagonistes. Il n'est pas possible de généraliser à partir des traits différenciateurs de la synonymie pour définir cette notion, puisqu'ils sont différents d'un couple de synonymes à l'autre. Si la synonymie est assimilée à une équivalence sémantique approchée, elle est susceptible d'englober les autres relations sémantiques. On propose donc de renouveler la description de l'équivalence sémantique en la décomposant en degrés d'intersection sémantique.

¹¹ Les sept péchés capitaux sont l'avarice, la colère, l'envie, la gourmandise, la luxure, l'orgueil et la paresse.

¹² Les quatre vertus cardinales sont la justice, le courage, la prudence et la tempérance ; les trois vertus théologales sont la charité, l'espérance et la foi.

¹³ C'est nous qui soulignons les lexicalisations étudiées dans chaque extrait du corpus.

¹⁴ Toutefois, l'enrichissement de la base lexicale ne peut se faire automatiquement à partir de la collecte des corrélats, tant rares sont ceux qui sont susceptibles de devenir des lexicalisations des thèmes.

¹⁵ Manifestent-elles les dernières traces des binômes synonymiques, même si ceux-ci ont disparu au début du XVII^e siècle (Buridant, 1980) ? Pour le vérifier, il faudrait confronter ces binômes à des textes antérieurs, reconnus pour les utiliser. Toutefois, les binômes sont seulement possibles dans des parataxes alliant deux lexicalisations ; dès lors qu'il y a trois lexicalisations, la structure binaire est rompue.

¹⁶ Il faut garder à l'esprit que les faits stylistiques demeurent souvent ponctuels et ne peuvent donc pas donner lieu à des phénomènes linguistiques récurrents comme ceux observés en linguistique de corpus.