



UNIVERSITÀ  
di VERONA

Dipartimento  
di LINGUE  
E LETTERATURE STRANIERE



Colloque International

# PhraséoTerm

## Phraséologie et Terminologie

Vérone, 16 - 17 novembre 2020

# Livret des résumés

Avec le parrainage de :  
  
**MINISTÈRE  
DE LA CULTURE** /  
*Liberté  
Égalité  
Fraternité*



[dls.univr.it](http://dls.univr.it) / [dh.dls.univr.it](http://dh.dls.univr.it)

## *Sommaire*

### **Conférences plénières**

Alain Polguère (Université de Lorraine, CNRS, ATILF), *Les modificateurs redondants et leur rôle en terminologie : de l'amour physique aux masques de protection*.....p. 5

Maria Teresa Zanola (Università Cattolica del Sacro Cuore, Milano), *La formation de phrasèmes terminologiques entre diachronie et variation*.....p. 6

### **Communications**

Silvia Calvi (Università di Verona), Patrick Drouin (Université de Montréal), Paolo Frassi (Università di Verona), *Acquisition automatique des collocations terminologiques : exploration dans le domaine du droit du commerce international*.....p. 8

Chen Ping-Hsueh (LIDILEM, Université Grenoble Alpes), Yan Rui (LIDILEM, Université Grenoble Alpes), *Les constructions Verbe causatif + Nom dans les textes institutionnels : approche contrastive (français/chinois)*.....p. 9

Valérie Delavigne (Clesthia EA 7345, Université Sorbonne Nouvelle), *Terme, phraséologie, discours : une approche socioterminologique*.....p. 11

Nataliia Demchuk (Université Ivan Franko - Lviv), *Analyse structurelle, sémantique et syntaxique des collocations dans le français du marketing*.....p. 12

Kaja Gostkowska (Université de Wrocław), *Les termes de couleur et leur entourage syntaxique : un champ lexical apte à la spécialisation ?*.....p. 14

Polina Mikhel (ATILF-CNRS, Université de Lorraine), Tomara Gotkova (ATILF-CNRS, Université de Lorraine), Francesca Inghrosso (LPCT-CNRS, Université de Lorraine), *Vocabulaire de la chimie : l'analyse sémantique des phrasèmes*.....p. 15

Galina Ovchinnikova (Université d'État pour la région de Moscou), Assya Ovchinnikova (Académie Internationale de Police, Toula), *Les moyens d'enrichissement du vocabulaire des coloronymes cosmétiques français et leurs équivalents russes et italiens*.....p. 16

Éric Poirier (Université du Québec à Trois Rivières), *Méthode contrastive et multifactorielle de repérage automatisé des expressions multilexémiques lexicalisées*.....p. 18

Chiara Preite (Università di Modena e Reggio Emilia), Daniela Dincă (Universitatea din Craiova), *Présence et traitement des phrasèmes dans les vocabulaires juridiques français*.....p. 20

Karina Rekosh (Institut de Moscou des Relations Internationales), *Phénomènes phraséologiques dans le français juridique*.....p. 21

- Delphine-Anne Rousseau (CRTT, Université Lumière Lyon 2), *Phrasèmes et collocations en langue de spécialité – le cas de la terminologie musicale ancienne*.....p. 24
- Michela Tonti (Università di Bologna DIT - Forlì), Pietro Moretti, « *Ajustement cosmétique* », « *don manuel* », « *rescrit social* » et bien d'autres : le phraséoterme à la confluence de la langue naturelle, de la langue de spécialité et des néoformations.....p. 25
- Agnès Tutin (Université Grenoble Alpes), *Quels traitements sémantiques pour la phraséologie scientifique transdisciplinaire dans une perspective d'aide à la rédaction ?*.....p. 26
- Federica Vezzani (Università degli Studi di Padova), *Vers une méthodologie pour l'extraction automatique des collocations en terminologie médicale*.....p. 27

## *Conférences plénières*

## Les modificateurs redondants et leur rôle en terminologie : de l'amour physique aux masques de protection

Alain Polguère (Université de Lorraine, CNRS, ATILF)

Le lexique du français, de l'anglais et, peut-être, de toute langue contient un nombre significatif de syntagmes nominaux phraséologiques « Nom + Modificateur » (ou « Modificateur + Nom ») où le Modificateur est sémantiquement redondant vis-à-vis du Nom. Ce Modificateur a pour fonction de marquer l'acceptation du vocable nominal impliquée dans le syntagme en question ; par exemple : *amour physique, communication téléphonique, compétition sportive, masque de protection, prise électrique...* Notons que, plus marginalement, on trouve les modificateurs redondants utilisés avec d'autres parties du discours ; par exemple, les Verbes – *aimer d'amour, désirer physiquement* – ou les Adjectifs – *physiquement fort*. De plus, ce type de redondance peut aussi se manifester sur le plan morphologique (Koch 2008 : 125-127).

Deux conditions doivent être satisfaites pour considérer qu'une construction « Nom + Modificateur » est un syntagme à modificateur redondant. D'une part, il faut diagnostiquer que ce syntagme n'est pas une locution – cf. Pereira de Abreu (2018), ainsi que la distinction formelle établie entre *dé à coudre* et *dé à jouer* dans Polguère (2011). D'autre part, il faut démontrer que la tête de ce syntagme est bien une acception spécifique du vocable correspondant. Ainsi, on peut démontrer qu'il existe une acception du vocable COMMUNICATION qui contient dans sa définition le sens 'téléphonique' et qui possède notamment ses collocatifs propres, marqueurs de son identité lexicale : *On a coupé la communication ; Attends, je prends la communication ;* etc. De ce fait, *communication téléphonique* est bel et bien un syntagme à modificateur redondant.

Les syntagmes à modificateur redondant, même s'ils sont par définition sémantiquement compositionnels (ce ne sont pas des locutions), sont clairement phraséologiques. Il s'agit de collocations, au sens de Hausmann (1979), qui ont pour base le gouverneur syntaxique (nominal, verbal...) et pour collocatif le modificateur redondant. Leur existence a été clairement reconnue dans le cadre de la Lexicologie Explicative et Combinatoire par l'introduction d'une fonction lexicale syntagmatique spécifique – **Redun** – dans le système des fonctions lexicales standards (Mel'čuk et Polguère, à paraître). Il se trouve que ce phénomène occupe une place non triviale dans la phraséologie des terminologies, où de nombreux termes scientifiques ou techniques dits *complexes* sont formellement des syntagmes nominaux à modificateur redondant : *bilan comptable, escompte de règlement, couple / fonction mathématique, genre grammatical, régime syntaxique, masque de protection*, etc. Dans bien des cas, la nature de syntagme redondant ne se manifeste que lorsque l'on confronte la nomenclature scientifique « théorique » à la réalité linguistique telle qu'elle se manifeste dans les textes écrits et oraux. Ainsi, s'il est vrai qu'il existe plusieurs types d'infarctus et que le syntagme *infarctus du myocarde* dénote une pathologie particulière parmi ceux-ci, dans la pratique, une acception individualisée de *infarctus* = lang. gén. *crise cardiaque* a émergé, qui désigne spécifiquement l'infarctus dit *du myocarde*.

Nous commencerons notre exposé par une modélisation lexicologique du phénomène général de la modification redondante : caractéristiques formelles, modélisation lexicographique et motivations sémiotiques. Nous aborderons ensuite le cas spécifique de la place des syntagmes nominaux à modificateur redondant dans les nomenclatures terminologiques et dans les textes (non) spécialisés.

### Références bibliographiques

- Hausmann, F. J. (1979). Un dictionnaire des collocations est-il possible ? *Travaux de littérature et de linguistique de l'Université de Strasbourg*, XVII(1), 187-195.
- Koch, P. (2008). Cognitive onomasiology and lexical change. Around the eye. In M. Vanhove (dir.) : *From Polysemy to Semantic Change*. Coll. « Studies in Language Companion Series (SLCS) » 106, Amsterdam/Philadelphie : John Benjamins, 107-137.
- Mel'čuk, I., Polguère, A. (à paraître). Les fonctions lexicales dernier cri. In S. Marengo (dir.) : *La Théorie Sens-Texte et ses applications. Lexicologie, lexicographie, terminologie, didactique des langues*. Coll. « Dixit Grammatica », Paris : L'Harmattan.
- Pereira de Abreu, S. (2018). Considerações sobre a natureza colocacional e locucional de sintagmas terminológicos. *ReDILLeT 1* [En ligne : <https://revistas.unc.edu.ar/index.php/ReDILLeT/article/view/22266/21866>].
- Polguère, A. (2011). Figement et ellipse dans une perspective lexicographique : le cas de *dé à jouer* et *dé à coudre*. In J.-C. Anscombe et S. Mejri (dir.) : *Le figement linguistique : la parole entravée*. Coll. « Lexica. Mots et Dictionnaires » 18, Paris : Champion, 363-373.

## La formation de phrasèmes terminologiques entre diachronie et variation

Maria Teresa Zanola (Università Cattolica del Sacro Cuore, Milano)

Pour établir une typologie de phrasèmes terminologiques en perspective diachronique, il faut défricher un terrain métaterminologique vaste, qui aide à s'orienter dans le chemin et à distinguer entre les archaïsmes et les phrasèmes qui restent vivants et/ou productifs dans le temps. La phraséologisation exprimée par une langue à une époque donnée et dans un domaine conceptuel déterminé – qui exprime l'état des connaissances et des coordonnées culturelles de la tranche temporelle considérée – montre la possibilité d'appliquer une méthode d'analyse solide et de pouvoir classer des phrasèmes sémantiques et non sémantiques (selon la catégorisation de Mel'čuk 1995 et 1997) en établissant différents degrés de figement.

L'efficacité de ces instruments d'analyse phraséologique sera illustrée à l'aide de quelques cas, entre diachronie et variation, au cours des siècles XV<sup>e</sup> et XVIII<sup>e</sup>. Dans le domaine des mathématiques, le *Triparty en la science des nombres* de Nicolas Chuquet (1484), le premier ouvrage d'algèbre écrit en français, offre un domaine très intéressant à observer pour les phrasèmes que l'on peut y repérer.

La valeur de la monnaie - moyen de paiement et de compte – est un terrain d'enquête aussi bien fructueux pour observer le figement : dans le recueil de problèmes arithmétiques du *Triparty*, les termes désignant les monnaies, tels que *denier*, *écu* et *marc d'or*, donnent lieu à des unités terminologiques figées sur lesquelles sera portée l'attention. L'analyse se déplace ensuite sur des phrasèmes terminologiques du XVIII<sup>e</sup> siècle, sur leur évolution et variation, à partir de types de textes différents et de domaines de la connaissances moins structurés par rapport au domaine scientifique.

Des exploitations d'ordre très diversifié permettront de mettre en valeur l'efficacité d'un classement théorique des phrasèmes terminologiques, qui permet de mettre en valeur les degrés de gradation - et de variation - que l'on peut établir dans le figement.

### Références bibliographiques

- Anscombe J.-C., Mejri S. 2011. *Le figement linguistique : la parole entravée*. Paris : Honoré Champion.
- Burger H., Dobrovolskij D., Kühn P. et Norrick N. (éd.). 2007. *Phraseology : An International Handbook of Contemporary Research*. Berlin-New York : Mouton de Gruyter.
- Cowie A. (éd.). 1998. *Phraseology : Theory, Analysis, and Application*. Oxford : Clarendon Press.
- Di Stefano G. 1991. *Dictionnaire des locutions en Moyen Français*. Montréal : Ceres.
- Fiala P., Lafin P. et Piguet M.-F. (éd.). 1997. *Locution : entre lexicologie, syntaxe et pragmatique*. Paris : Klincksieck.
- Mel'čuk I.A. 1995. « Phrasemes in Language and Phraseology in Linguistics ». In M. Everaert, E.-J. van der Linden, A. Schenk et R. Schreuder (éd.), *Idioms. Structural and Psychological Perspectives*, Hillsdale, N.J. – Hove : Erlbaum, 167-232.
- Mel'čuk I.A. 1997. *Cours de morphologie*, vol. 4. Montréal : Presses de l'Université de Montréal.
- Mel'čuk I.A., Clas A., Polguère A. 1995. *Introduction à la lexicologie explicative et combinatoire*. Bruxelles : Duculot.
- Zanola M.T. 2003. « Phrasèmes sémantiques et pragmatiques dans le lexique des sciences du XV<sup>e</sup> siècle : pour une typologie phraséologique ». *Le Moyen Français*, 51-52-53, 641-650.

## *Communications*

## Acquisition automatique des collocations terminologiques : exploration dans le domaine du droit du commerce international

Silvia Calvi (Università di Verona)  
Patrick Drouin (Université de Montréal)  
Paolo Frassi (Università di Verona)

L'extraction terminologique constitue un domaine d'investigation de grande importance du Traitement automatique de la langue (TAL) : linguistes, terminologues et informaticiens travaillent au développement de logiciels capables d'automatiser autant que possible l'acquisition automatique des termes. Même si l'extraction des termes est de mieux en mieux prise en charge par les techniques informatiques, ce n'est pas le cas de toutes les entités terminologiques qui caractérisent la langue spécialisée.

Dans le cadre de cette communication, nous proposons d'analyser un type particulier d'unité phraséologique, la *collocation*. Cette dernière est un phrasème lexical semi-contraint composé de deux éléments dont l'un (la *base*) est choisi librement par le locuteur, tandis que l'autre (le *collocatif*) est contraint en fonction du sens que le locuteur veut exprimer (Mel'čuk 2013). Plus particulièrement nous proposons d'investiguer la question de l'extraction automatique de cette unité phraséologique en langue de spécialité. En effet, les propriétés linguistiques des collocations posent des problèmes concernant leur extraction automatique. Par exemple, le fait que ces unités phraséologiques peuvent se présenter de manière disjointe (*la sentence, exécutoire, s'impose ; atteindre le plus grand nombre possible de clients*) rend plus difficile leur extraction par rapport à d'autres unités phraséologiques, comme, par exemple, les locutions.

Notre objectif principal est de décrire une méthode d'acquisition semi-automatique des collocations terminologiques qui dans nos essais pratiques a donné les résultats les plus performants. Ces essais ont été réalisés à partir d'un corpus de droit du commerce international élaboré par l'équipe de recherche DiaCom<sup>1</sup> du département des Langues et Littératures étrangères de l'Université de Vérone. Notre but ultime est de décrire les collocations de ce domaine dans une base de données terminologiques de type *réseau lexical* (Polguère 2014).

D'abord, nous présenterons les caractéristiques syntaxiques et sémantiques de la collocation qui la distinguent des autres unités phraséologiques. Nous avons adopté la définition proposée par la *Lexicologie Explicative et Combinatoire* (Mel'čuk 1993, 2013 ; Wanner 1996) qui, pouvant s'appliquer à la langue générale comme à la langue de spécialité (Frassi *et al.* 2020), offre une classification détaillée que nous voulons reproduire dans notre base de données terminologiques. Après avoir passé en revue les travaux les plus significatifs sur l'extraction automatique des collocations en langue générale et en langue de spécialité (Smadja 1993 ; Williams 1998 ; Maniez 2002 ; Orliac 2006 ; Delač *et al.* 2009), nous donnerons des exemples d'outils permettant de prendre en charge l'extraction des collocations. Nous avons choisi trois outils aux fonctionnalités différentes – le concordancier *AntConc* (Anthony 2019), l'extracteur des termes *TermoStat* (Drouin 2003) et le gestionnaire de corpus *Sketch Engine* (Kilgarriff *et al.* 2014) – pour comparer les résultats obtenus et en comprendre les points forts et les faiblesses. Plus précisément, nous analyserons les fonctions de ces logiciels qui permettent d'extraire des collocations : *Collocate* de *AntConc*, *Bigrammes* de *TermoStat* et *Word Sketch* de *Sketch Engine*. Nous présenterons des essais d'extraction automatique des collocations terminologiques à partir de notre corpus du droit du commerce international : un corpus en langue française composé de 37 textes, 830 355 occurrences au total. Notre intérêt se portera sur plusieurs types de collocations que nous classerons selon une approche sémantique, en ayant recours au concept de *Fonctions Lexicales* développé dans le cadre de la *LEC* (Wanner 1996), et syntaxique, en étudiant les patrons syntaxiques les plus récurrents, soit nom+ verbe, verbe+ nom, adjectif+ nom, nom+ adjectif, nom+ préposition+ nom, verbe+ adverbe.

Les premiers essais d'extraction automatique nous feront comprendre les failles des outils présentés. Par exemple, la fonction *Collocate* de *AntConc* conduit à des combinaisons de mots qui, bien souvent, ne correspondent pas à la définition de *collocation* que nous avons adoptée : en effet il s'agit des combinaisons récurrentes d'un point de vue quantitatif mais qui ne présentent pas les propriétés syntactico-sémantiques des collocations que nous avons illustrées (*sentence arbitrale, sentence et sa validité*). Pour sa part, *TermoStat* donne comme termes complexes seulement des syntagmes nominaux (*clause compromissoire, droit international*) qui, dans la plupart des cas, ne sont pas de collocations. La fonction *Bigrammes* du même logiciel permet d'analyser les collocations à collocatif verbal (*remettre en cause, apporter une solution*), fonction de grand intérêt si nous considérons que l'acquisition de cette typologie de collocation crée encore aujourd'hui de nombreuses difficultés (souvent liées au fait qu'elles peuvent se présenter de manière disjointe). La fonction *Word Sketch* de *Sketch Engine* présente aussi une faille : elle ne permet pas de différencier la collocation des autres types d'unités phraséologiques, plus particulièrement des locutions, comme (*commerce*) *en ligne* et *commerce électronique*.

Ces difficultés nous obligeront à adopter d'autres méthodes d'extraction semi-automatique. Nous proposerons de comparer les résultats obtenus par *AntConc*, *TermoStat* et *Sketch Engine* : nous créerons une première liste comprenant

---

<sup>1</sup> Le sous-corpus du domaine du droit du commerce international provient d'un plus grand corpus en voie de constitution à l'Université de Vérone. La réalisation d'un corpus de commerce international par notre équipe de recherche DiaCom s'insère dans le cadre du projet d'excellence *Les humanités numériques appliquées aux langues et aux littératures étrangères* pour lequel l'université de Vérone a obtenu un financement de la part du Ministère italien de l'éducation, de l'université et de la recherche.



des *candidats de collocations* extraits automatiquement. À partir de cette liste, nous repérerons manuellement les unités phraséologiques de type *collocation*. Nous analyserons les *collocations candidates* selon trois critères : 1) appartiennent-elles au domaine d'intérêt ? 2) présentent-elles un des patrons syntaxiques que nous avons illustrés ? 3) peuvent-elles se décrire en faisant appel aux fonctions lexicales introduites par la *LEC* ?

Cette méthode permettra de réaliser une liste d'une centaine de collocations classées selon leur patron syntaxique et leur sémantisme, représentant un échantillon important du domaine, tant du point de vue qualitatif que du point de vue quantitatif.

### Références bibliographiques

- Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan : Waseda University. Disponible sur <https://www.laurenceanthony.net/software>
- Delač, D. *et al.* (2009, March). TermeX : a tool for collocation extraction. In *International Conference on Intelligent Text Processing and Computational Linguistics*, 149-157, Springer.
- Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, 9(1), 99-115.
- Frassi, P. *et al.* (2020). Fouille de textes et repérage d'unités phraséologiques. In Brun, C., Roche, C. (éds.) *TOTh 2019 Terminologie & Ontologie*, 321-338.
- Kilgarriff, A. *et al.* (2014). The Sketch Engine : ten years on. *Lexicography*, 1, 7-36.
- Maniez, F. (2002). Un modèle d'extraction des collocations en langue de spécialité. *ASp. la revue du GERAS*, 35-36, 35-47.
- Mel'čuk, I. (1993). La Phraséologie et son rôle dans l'enseignement-apprentissage d'une langue étrangère. *Études de linguistique appliquée*, 92, 82-113.
- Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais. *Cahiers de lexicologie*, 102(1), 129-149.
- Orliac, B. (2006). Colex : un outil d'extraction de collocations spécialisées basé sur les fonctions lexicales. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 12(2), 261-280.
- Polguère, A. (2014). From writing dictionaries to weaving lexical networks. *International Journal of Lexicography*, 27(4), 396-418.
- Smadja, F. (1993). Retrieving collocations from text : Xtract. *Computational linguistics*, 19(1), 143-177.
- Wanner, L. (Ed.). (1996). *Lexical functions in lexicography and natural language processing*, 31, John Benjamins Publishing.
- Williams, G. (1998). Collocational networks : Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, 3(1), 151-171.

## Les constructions *Verbe causatif* + *Nom* dans les textes institutionnels : approche contrastive (français/chinois)

Chen Ping-Hsueh (LIDILEM, Université Grenoble Alpes)  
Yan Rui (LIDILEM, Université Grenoble Alpes)

Dans cette étude, nous proposons de comparer les combinaisons *Verbe causatif* + *Nom* avec des structures ayant les mêmes sens en chinois dans le corpus institutionnel de l'*Organisation des Nations Unies* (désormais, *ONU*)<sup>1</sup>. L'intérêt de notre étude se justifie d'abord parce qu'il est intéressant de comparer le français avec une langue distante comme le chinois pour voir les différents moyens utilisés dans l'expression de la causalité ; et ensuite parce que ces collocations lexicales définies comme des associations (semi-)figées et sémantiquement compositionnelles (Tutin 2013) sont fréquentes en langue de spécialité et constituent souvent des déficits persistants des apprenants du FLE (Bak Sienkiewicz 2016).

En nous basant sur le classement des mécanismes causatifs de Dixon (2000) qui catégorise les moyens d'expression de la cause en fonction de leur compacité morpho-syntaxique, (par exemple, les verbes causatifs *causer*, *provoquer* sont les plus compacts), nous cherchons à mettre en évidence les relations syntaxico-sémantiques entre les verbes causatifs et les noms dans un corpus parallèle bilingue (français→chinois) composé de textes institutionnels de l'Organisation des Nations Unies. Le corpus français est de 585,015,613 mots alors que le corpus chinois comporte 243,376,079 caractères. D'un point de vue méthodologique, une liste de verbes causatifs français a été établie au préalable à partir de travaux antérieurs sur la causalité (Gross *et al.* 2009 ; Diwersy et François 2011). Vu la taille importante du corpus, le seuil de fréquence a été fixé à 1000 occurrences. Nous avons obtenu par la suite une liste de 32 verbes causatifs français à étudier. Le tableau 1 illustre les 5 verbes causatifs les plus récurrents et les 3 arguments les plus fréquents selon l'indice « LogDice »<sup>2</sup> :

Verbes causatifs (dimension sémantique)	Arguments	Score (LogDice)
renforcer (intensité forte)	capacité	12,17
	coopération	10,97

<sup>1</sup> Ce corpus est accessible sur la plateforme *Sketch Engine* (<https://www.sketchengine.eu/>).

<sup>2</sup> L'indice « LogDice » est une valeur de significativité des cooccurrences des lexies, utilisée dans *Sketch Engine*. D'ailleurs, le score peut atteindre jusqu'à 14. Cela signifie que toutes les occurrences du mot A se produisent avec celles du mot B et vice-versa.

	rôle	9,28
réduire (intensité faible)	pauvreté	10,8
	émission	10,48
	coût	10,2
éliminer (aspectuel terminatif)	pauvreté	11,37
	terrorisme	10,74
	discrimination	10,69
accroître (intensité forte)	participation	9,74
	transparence	9,73
	nombre	9,68
entraîner (neutre)	augmentation	10,01
	mort	10,01
	réduction	9,39

Tableau 1. Les 5 verbes causatifs français les plus fréquents et leurs arguments.

Il est intéressant de constater que les verbes causatifs d'intensité forte (*accroître, renforcer*) ont des arguments plus ou moins positifs ou neutres. Tandis que les verbes causatifs d'intensité faible ou aspectuels terminatifs (*éliminer, réduire*) s'associent plus souvent avec des noms de polarité négative. D'ailleurs, l'attraction entre le verbe causatif et le nom est très forte. En revanche, la polarité de la combinaison *Verbe causatif + Nom* est relativement positive, telle que *renforcer + capacité ; accroître + transparence ; réduire + pauvreté ; éliminer + discrimination*. Ce constat peut être lié au genre textuel du corpus. En effet, l'ONU a pour objectif de promouvoir la paix et les liens amicaux entre les nations. D'un point de vue contrastif, en comparant ces résultats avec le chinois, nous avons remarqué que ce dernier dispose de 3 moyens principaux pour traduire l'association de *Vcaus + N*.

1. V causatif chinois :

(1) [...] mais beaucoup de ces conflits **ont** [...] **entraîné des millions de morts**. (*Sketch Engine, MultiUN*)

[...] 但许多/冲突 [...] **造成/数百万/人/死亡**。

[...] dàn / xǔduō / chōngtú [...] **zàochéng / shù / bǎiwàn / rén / sǐwáng**.

Litt. [...] mais / beaucoup / conflits [...] **causer / quelque / million / personne / mort**

2. V support de causation + V2/adj. :

(2) Soulignant la nécessité de **renforcer la capacité** de prévention des conflits et de maintien de la paix en Afrique, [...]. (*Sketch Engine, MultiUN*)

强调/必须**加强**/非洲/预防/冲突/和/维持/和平的/能力, [...].

qiángdiào / bìxū / **jiāqiáng** / fēizhōu / yùfáng / chōngtú / hé / wéichí / héping / de / **nénglì**, [...].

Litt. Souligner / devoir / **V support de causation + fort (adj.)** / Afrique / prévenir / conflit / et / maintenir / paix / DE subordonateur nominal / **capacité**, [...]

3. Périphrase causative V1 causatif + V2 non causatif :

(3) Outre l'objectif qui consiste à **accroître la participation** directe et la responsabilité des Timorais dans le domaine politique, [...]. (*Sketch Engine, MultiUN*)

除/在/政治/领域/使/帝汶人/更多/地/直接/参与/和/承担/责任/这个/目标/外, [...].

chú / zài / zhèngzhì / língyù / **shǐ** / Dìwènrén / gèng duō / de / zhíjiē / **cānyù** / hé / chéngdān / zérèn / zhè / gè / mùbiāo / wài, [...].

Litt. Outre / dans / politique / domaine / **faire en sorte que** / Timorais / davantage / DI particule d'adverbalisation / directement / **participer** / et / assurer / responsabilité / ce / CL / objectif / extérieur, [...]

D'abord, il existe des verbes causatifs chinois comme équivalents, comme dans l'exemple 1 (造成 *zàochéng*). Ensuite, il est intéressant de constater que dans l'exemple 2 l'ajout du verbe support de causation (加 *jiā*) entraîne la transitivation et la causativisation du V2/adj (强 *qiáng*) (Lin 2001 ; Basciano 2010). Enfin, nos résultats rejoignent l'étude de Basciano (2010, 121-123) selon laquelle les périphrases causatives chinoises sont un des moyens les plus souvent employés pour exprimer la causalité en chinois moderne et contemporain.

A l'issue de l'étude comparative, nous proposons un éventail des équivalents fonctionnels chinois des combinaisons *verbe causatif français + Nom*. Cet éventail permet de mieux appréhender comment ces constructions sont traduites en chinois.

## Références bibliographiques

- Bak Sienkiewicz, M. (2016). *Les constructions Verbe causatif + Nom d'émotion. Aspects linguistiques et pistes didactiques* (Thèse de doctorat). Université Grenoble Alpes.
- Basciano, B. (2010). *Verbal compounding and causativity in Mandarin Chinese* (Thèse de doctorat). Università di Verona.
- Chen, P. H. 陳品學. (2017). *Les verbes causatifs français et leurs équivalents en chinois (étude contrastive)* (Mémoire de master 2 Linguistique non publié). Université Grenoble Alpes.

- Diwersy, S. et François, J. (2011). La combinatoire des noms d'affect et des verbes supports de causation en français. Étude de leur attirance au niveau des unités et de leurs classes syntactico-sémantiques. *Revue TRANEL (Travaux neuchâtelois de linguistique)*, 55, 139-161.
- Dixon, R.M.W. (2000). A typology of causatives : form, syntax and meaning. In Dixon R.M.W. and A. Aikhenvald (éds.), *Changing valency. Case studies in transitivity*, 30-83. Cambridge University Press.
- Gross, G., Pauna, R. et Valetopoulos, F. (2009). *Sémantique de la cause*. Peeters.
- Lin, T.-H. 林宗宏. (2001). *Light Verb Syntax and the Theory of Phrase Structure* (Thèse de doctorat). University of California.
- Tutin, A. (2013). Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument, *Langages*, 189 (1), 47-63.

## Terme, phraséologie, discours : une approche socioterminologique

Valérie Delavigne (Clesthia EA 7345, Université Sorbonne Nouvelle)

Les discours spécialisés se caractérisent par la mise en œuvre de terminologies spécifiques à une pratique sociale située. Ces discours se distinguent par l'usage de termes, unités lexicales à la définition stabilisée dans des communautés discursives données, au mode de formation particulier et au statut sémiotique spécifié par les normes en vigueur (ISO 704 et 1087)<sup>1</sup>. Si l'on envisage les termes dans leurs fonctionnements langagiers, la problématique du « mot juste » (Mortureux 1994) se trouve reliée de facto à la question des genres discursifs, des situations, voire des thèmes abordés. Analyser le fonctionnement des termes invite à les mettre en relation avec la diversité des discours dans lesquels ils apparaissent et leurs différents modes de circulation, de réception et d'interprétation. Marqueurs de connaissance, marqueurs de savoir-faire, marqueurs identitaires : la fonction sociolinguistique des termes, loin d'être à négliger, conditionne leurs conditions d'apparition. Elle pousse à ne plus opposer la « langue générale », bien rarement définie, à des « langues de spécialité » pour repenser le système dans sa globalité en fonction des usages et donc, des discours, ce qu'a mis en évidence notamment la socioterminologie (Gaudin 2003).

Au-delà des termes, les discours spécialisés actualisent divers types de solidarités lexicales récurrentes, dont le statut n'est pas toujours clair. Si, à l'aube des années 2000, on pouvait faire le constat d'un manque d'études sur la phraséologie, les bibliographies sur le thème sont aujourd'hui copieuses. Depuis plusieurs années, nombre d'études linguistiques se penchent sur le préfabriqué dans la langue, le « déjà-là », « prêt à parler » ou « prêt à écrire », « motifs », « routines » et « autres « formules », séquences plus ou moins figées impliquant divers niveaux linguistiques qui sollicitent à la fois le lexique, la syntaxe, la sémantique, dont l'appel au colloque se fait l'écho (Krieg-Planque 2009 ; Legallois et Tutin 2013 ; Misri 1987 ; Née *et al.* 2016 ; Plane et Rondelli 2017 ; Sitri et Tutin 2016). Les discours spécialisés ne dérogent pas aux règles de la langue : des séquences de nature variable, plus ou moins figées, y sont non seulement présentes, mais, dans certaines sphères d'activité, sont bien plus fréquentes que dans d'autres types de discours (Aussenac-Gilles et Condamines 2009 ; Gautier 2004, 2018 ; Pecman 2018). On y trouve aux côtés de termes complexes – parfois d'une longueur telle que la question de leur découpage se pose et que l'on peut se demander si on doit les identifier comme une ou plusieurs unités (Assal et Delavigne 1993) – des collocations et autres séquences phraséologiques, terminologiques ou non d'ailleurs, qui permettent d'instancier les termes. Certains aspects syntagmatiques relèvent-ils de la phraséologie — que l'on peut définir rapidement comme des combinaisons caractéristiques d'un discours donné — ou comme des unités terminologiques uniques ? Dans l'énoncé « on parle de ganglion accolé », à la structure métadiscursive caractéristique du pointage vers une unité terminologique, *ganglion accolé* peut être considéré comme une collocation, un terme à part entière ou une unité phraséologique selon les critères que l'on se donne<sup>2</sup>. Quel que soit ce statut, si on considère que le rôle essentiel des terminologies est leur fonctionnalité dans une communauté discursive donnée, ces associations privilégiées doivent être repérées et décrites.

Ces questions sont d'importance dès lors que l'on souhaite constituer une ressource terminographique quelle qu'elle soit, base de données, dictionnaire ou glossaire. Comment les repérer ? Sur quels critères les retenir ? Quelle place leur donner ? Comment les décrire ? Comment intégrer les variations repérées ? Car ce n'est plus seulement le « terme juste » qu'il s'agit de pointer, mais sa mise en discours qu'il convient de faciliter.

Outre les aspects théoriques que ces questions engagent, elles se posent à nous dans le cadre de deux projets, l'un autour de « l'expertise » de patients atteints de cancer, l'autre à propos d'un ouvrage lexicographique autour de l'écologie. Dans des communications élargies qui dépassent les communautés discursives où les termes circulent d'ordinaire, ces formations préférentielles peuvent, à l'instar des termes, devenir un écueil : pour savoir utiliser le terme *hormone*, il peut être utile de savoir qu'elle doit être *secrétée*, qu'une *perfusion* doit être *posée*, que des *produits radioactifs émettent des rayonnements*, ou qu'un *traitement* doit *faire la preuve de son efficacité*.

Nous nous proposons d'articuler notre propos de la façon suivante : après une rapide présentation des enjeux des projets et des corpus sollicités, nous souhaiterions revenir sur la définition de ces combinaisons privilégiées et la question des méthodologies à mettre en œuvre pour leur repérage, notamment en tenant compte des variations d'usage. Nous

<sup>1</sup> Ce statut sémiotique est partiellement validé selon les corpus (Gaudin 2003 ; Petit 2018).

<sup>2</sup> Tout comme « on parle de » qui, à l'instar de la phraséologie scientifique, peut être considéré comme routines des discours de vulgarisation (Delavigne 2020 ; Tutin et Grossmann 2014).

discuterons ensuite de la façon dont ces formes peuvent intégrer une base de données pour non spécialistes, élaborée dans une perspective « d'équipement » terminologique et phraséologique.

## Références bibliographiques

- Assal, A., Delavigne, V. (1993). Le découpage des unités terminologiques complexes : limites des critères linguistiques. *Les langues de spécialité : pratiques, outils, théories. Actes de la 4e Journée ERLA-GLAT*, 175-193.
- Aussenac-Gilles, N., Condamines, A. (2009). Variation syntaxique et contextuelle dans la mise au point de patrons de relations sémantiques. In Minel J.-L. (éd.), *Filtrage sémantique*, Paris : Hermès/Lavoisier, 115-149.
- Delavigne, V. (2020). Phraséologie et didacticité dans les discours de vulgarisation médicale : une ergonomie discursive. In Albano M., Savatovsky D. (éds.), *Phrasis. Repenser le figement : enjeux et perspectives contrastives en phraséodidactique des langues*.
- Gaudin, F. (2003). *Socioterminologie : une approche sociolinguistique de la terminologie*. Bruxelles : De Boeck Duculot (Champs linguistiques. Manuels).
- Gautier, L. (2004). Terme, phraséotermes, phrasème : questions de délimitation en langue spécialisée. In Jatlaoui H., E.M. Salah Méri (éds.), *Le Continuum en linguistique*, Sousse, Tunisia, Faculté des Lettres et Sciences Humaines de Sousse, 153-172.
- Gautier, L. (dir.) (2018). *Figement et discours spécialisés*, Berlin : Frank & Timme, Verlag für wissenschaftliche Literatur (Forum für Fachsprachen-Forschung).
- Krieg-Planque, A. (2009). *La notion de « formule » en analyse du discours : cadre théorique et méthodologique*, Besançon : Presses Universitaires de Franche-Comté (Annales littéraires de l'Université de Franche-Comté. Série Linguistique et sémiotique).
- Legallois, D., Tutin, A. (2013). Présentation. Vers une extension du domaine de la phraséologie. *Langages*, 189, 3-25.
- Misri, G. (1987). Approches du figement linguistique : critères et tendances. *La Linguistique*, vol. 23, n° 2, 71-85.
- Mortureux, M.-F. (1994). Comment définir la propriété d'un mot. In Moirand S., Abdelmadjid A.B., Beacco J.-C., Collinot A. (éds.), *Parcours linguistiques de discours spécialisés*, Berne ; New York : Peter Lang (Sciences pour la communication), 3-10.
- Née, É., Sitri, F., Veniard, M. (2016). Les routines, une catégorie pour l'analyse de discours : le cas des rapports éducatifs. *Lidil. Revue de linguistique et de didactique des langues*, n° 53, 71-93.
- Pecman, M. (2018). *Langue et construction de connaissances : énergie lexico-discursive et potentiel sémiotique des sciences*, Paris : L'Harmattan.
- Petit, G. (2018). Le figement : une asymptote pour la constitution des terminologies. In Gautier L. (éd.), *Figement et discours spécialisés*, Berlin : Frank & Timme, Verlag für wissenschaftliche Literatur (Forum für Fachsprachen-Forschung), 27-56.
- Plane, S., Rondelli, F. (2017). Le déjà-là dans l'écriture : quel substrat pour quels (ré)emplois ? *Pratiques. Linguistique, littérature, didactique*, n° 173-174.
- Sitri, F., Tutin, A. (2016). Présentation. *Lidil. Revue de linguistique et de didactique des langues*, n° 53, 5-18.
- Tutin, A., Grossmann, F. (éds.) (2014). *L'écrit scientifique : du lexique au discours : autour de Scientext*, Rennes : Presses universitaires de Rennes.

## Analyse structurelle, sémantique et syntaxique des collocations dans le français du marketing

Nataliia Demchuk (Université Ivan Franko - Lviv)

Les moyens d'expression de la communication spécialisée comprennent non seulement les termes simples ou composés, mais également diverses combinaisons de termes entre eux ou avec des mots de la langue commune : collocations (co-occurrences), phrasèmes et, dans un sens plus large, les locutions standard (Recommandations relatives à la terminologie, 2014). L'analyse des textes spécialisés a mis en évidence le rôle fondamental des termes pour ce type de discours et, au-delà, la nature phraséologique de la terminologie de marketing.

### 1. Collocations terminologiques

En règle générale, la nécessité de refléter plus en détail les caractéristiques d'un concept conduit à la formation des termes dont le nombre d'éléments dépend du degré de complexité de sa structure : plus le nombre est élevé, plus la signification du terme est détaillée, plus sa définition est précise.

#### 1.1 Création des collocations en langues de spécialités

Le collocateur en tant que porteur de la dénomination descriptive, peut rester invariable tandis que le collocatif acquiert la signification sémantique résultant de la création des collocations terminologiques (L'Homme 2004 : 113). Pour déterminer la genèse des collocations – *structures polylexicales, dont les éléments sont disposés d'une façon cohérente et dans un ordre sémiotique différent* – il en existe deux processus opposés : par restriction du sens initial dans lequel est pris un terme ou bien par expansion de sa forme externe, ce qui est nécessaire à la concrétisation de concept lui-même. Cela conduit à la fixation de nouveaux concepts ou phénomènes au moyen de la création des collocations terminologiques, ce qui assure la rigueur des liens au sein du système des termes trouvés dans le champ scientifique concerné :

- **informations (f pl) – informations primaires – informations primaires souhaitées – recueil d'informations primaires souhaitées ;**
- **entreprendre le recueil d'informations primaires souhaitées – entreprendre les méthodes du recueil d'informations primaires souhaitées – entreprendre les méthodes du recueil d'informations primaires souhaitées sur le marché.**

Une collocation présente donc des avantages par rapport au terme simple dans le sens où le terme devient plus spécifique avec l'augmentation de nombre de collocatifs qui l'intègrent : la base « informations » est liée par une affinité de nature notionnelle à des collocatifs par expansion en vue d'une description précise.

### 1.2 Structure des collocations en terminologie

Il conviendrait de dire que l'analyse du lexique de marketing a révélé la composition morphologique de collocations à deux (*prix fixé*), trois (*à prix fort*, *prix abusivement bas*, *niveau général des prix*), quatre (*prix toutes taxes comprises*), cinq éléments (*remboursement à terme au prix d'achat*), ce qui a permis de retracer le processus de complication de la structure des énoncés, c'est-à-dire observer la transformation des collocations simples à deux composants en une structure plus complexe, constituée de plusieurs éléments et d'identifier comment la structure influence leur volume. Comme on le voit dans les exemples ci-dessus, la base est le porteur du nom descriptif, tandis que les collocatifs précisent la spécialisation du champ de connaissances déterminé. En règle générale, les bases sont des prédicats nominaux empruntés au langage commun, et les collocatifs sont des adjectifs ayant la capacité d'élargir le sens d'un mot ou d'un terme, ce qui nécessite sans doute la connaissance du sujet étudié (González 2002 : 89-94).

Le principe d'un modèle structurel des combinaisons des noms, caractérisé en tant qu'une unité grammaticale du nom et de ses éléments associés, constitue un conditionnement catégoriel du mot de base ainsi que de ses collocatifs. En marketing sont surtout présentes les collocations terminologiques à la subordination cohérente des composants qui s'enchaînent. Un mot qui remplit la fonction d'un collocatif à un niveau, devient un mot de base à un autre niveau (*élasticité de la demande – la demande élastique ; bien de consommation durable – la durée de vie d'un produit*). En conséquence, chaque composant de la collocation est à un niveau de subordination inférieur par rapport au précédent.

### 1.3 Modèles de formation des collocations en marketing

Du point de vue grammatical dans le lexique de marketing nous avons établi la prévalence de deux modèles, composés d'au moins de deux mots : *N + Adj* : *emballage secondaire*, *position concurrentielle* et *N + prép. + N* : *fabrication en série*, *marché sans frontières*.

Les termes qui sont combinés avec le mot de base pour former une collocation, en constituent une unité sémantique.

### 2. Critère sémantique en tant que potentiel communicatif des collocations

Le critère sémantique affermit la signification des collocations. A. Lehmann, F. Martin-Berthet parlent des expressions, dans lesquelles les éléments rajoutés rendent plus individuelle l'unité lexicale formée, en séparant son contenu du contenu de ses parties (Lehmann, Martin-Berthet 2000 : 224). La liberté de choix des moyens linguistiques, afin de remplir les collocations d'une certaine valeur sémantique, est définie par certains cadres communicatifs et pragmatiques du destinataire : caractéristiques socioprofessionnelles, conditions et lieu de communication, etc. Un tel choix est soumis à des tâches communicatives spécifiques, ainsi qu'au principe de l'opportunité communicative. Les combinaisons de certains termes ne s'expliquent que par des subtilités sémantiques : par exemple, la formation de collocation *prix compétitifs* est due aux caractéristiques générales reconnues par les textes spécialisés, en particulier la neutralité, la précision, la concision (selon C. Delplanque). À notre avis, dans cette situation il serait logique d'utiliser l'adjectif *concurrentiel* (*prix concurrentiels*), car il ne s'agit pas de concurrence en tant qu'elle, mais de la concurrence entre monopoles, dont dépendent les prix intérieurs.

### 3. Collocations et unités phraséologiques

En règle générale, l'intégrité de la dénomination, le caractère fonctionnel stable de ses composants, sa reproductivité rapprochent les collocations des unités phraséologiques. Dans les collocations : *animation de l'unité commerciale*, *théâtralisation de l'offre*, *séduire le client*, *le confort d'achat*, le rôle des éléments métaphorisés (animation f, théâtralisation f, séduire v, confort m) est d'imposer psychologiquement la présentabilité d'un produit, ce qui se reflète dans le comportement du consommateur et augmente l'intérêt pour ce produit (l'expansion). Les linguistes (F. Hausmann, I. Mel'čuk, C. Deplanque) soulèvent le sujet de la phraséologie collocative dans leurs recherches en nous renvoyant au fonctionnement des constructions lexicales binaires dans le langage, dont la combinaison est basée sur la fonction sémantique (Mel'čuk, 1995).

#### 3.1 Rôle du contexte

Ce sont les caractéristiques structurelles et sémantiques qui déterminent le potentiel communicatif des unités phraséologiques, assurent leur formation, leur fonctionnement et leur perception par la communauté scientifique. Plusieurs linguistes évoquent la nécessité d'implication des facteurs contextuels dans l'analyse du sens des collocations. Il s'agit de prendre en compte dans une certaine mesure l'aspect pragmatique dans le discours associé, les particularités du langage de ses usagers. Ainsi, faire une analogie entre « la magie », en tant que l'art d'un illusionniste pour attirer l'attention du public, et les actions marketing, dont l'objectif est de vendre avec succès des biens, des services, etc., se reflète dans la fonction pragmatique dans les exemples pris de l'expansion : *La Banque asiatique de développement a fait un peu de magie comptable pour augmenter le montant de ses prêts ; ... le compte rendu des commerciaux chinois a fait l'effet d'une douche froide : les téléphones de Xiaomi, le géant chinois des smartphones, ne se vendent plus* (Courrier international). Dans le cas examiné l'effet « d'une douche froide » signifie un arrêt brutal des ventes, ce qui choque certainement l'acheteur. C'est à travers l'analyse de la structure communicative que les expressions linguistiques sont utilisées dans la terminologie professionnelle.

Pour conclure, les collocations terminologiques permettent de transmettre d'une manière adéquate et de refléter plus en détail les caractéristiques différentielles des concepts, ainsi que de révéler leur nature notionnelle. Leur structure et leur formation s'inscrivent dans les normes de la langue générale. Dans la terminologie les unités collocationnelles, ainsi que les unités phraséologiques sont plus résistantes que les collocations de la langue générale. Les collocations du point de

vue de leur structure ne sont pas créées, mais reproduites selon la forme préétablie en raison de l'économie des efforts et de temps. Ces unités linguistiques reproductrices signifient une combinaison de deux ou plusieurs mots ce qui les prédisposent à l'unification mutuelle par les rapports sémantiques et grammaticaux. Dans ce sens, le contexte est important car il démontre leur emploi dans les langues de spécialités et reflète plus clairement les spécificités de la pensée et de la vision du monde des membres d'une communauté linguistique concernée.

### Références bibliographiques

- Binon, J. *et al.* (2000). *Dictionnaire d'apprentissage du français des affaires*. Paris : Didier.
- Delplanque, C. (1999). Le statut phraséologique des composés économiques allemands. In *Les Cahiers du CIEL*. Paris : Université de Paris VII. Version électronique : <http://rech.eila.jussieu.fr/CIEL/cahiers/96-97/3Delplanque.html>.
- González Rey, M. I. (2002). *La phraséologie du français*. Toulouse : Presses Universitaires du Mirail.
- Hausmann, F. J. (1979). *Un dictionnaire des collocations est-il possible ?* In *TraLiLi*, 17, 187-195.
- Lehmann, A. *et al.* (2000). *Introduction à la lexicologie. Sémantique et morphologie*. Paris : Nathan.
- L'Homme, M.-C. (2008). *La terminologie : principes et techniques*. Montréal : Les Presses de l'Université de Montréal, 2008.
- Madry, P. (2009). *Dictionnaire pratique du commerce*. Bruxelles : De Boeck.
- Mel'čuk, I. *et al.* (1995). *Introduction à la lexicologie explicative et combinatoire*. Bruxelles : Duculot.
- Phraséologie spécialisée / Recommandations relatives à la terminologie* (entièrement révisée en collaboration avec la Section de terminologie de la Chancellerie fédérale suisse) : [http://www.cotsoes.org/sites/default/files/CST\\_Recommandations\\_relatives\\_a\\_la\\_terminologie\\_2014.pdf](http://www.cotsoes.org/sites/default/files/CST_Recommandations_relatives_a_la_terminologie_2014.pdf).

## Les termes de couleur et leur entourage syntaxique : un champ lexical apte à la spécialisation ?

Kaja Gostkowska (Université de Wrocław)

Pour notre communication, nous partons du point de vue de la terminologie variationniste, c'est-à-dire l'étude des variations de la langue spécialisée à plusieurs niveaux : « variation des langues, des discours et des textes de spécialité ; variation dans le temps, dans l'espace, en fonction des interlocuteurs et des situations de communication ; variation des unités linguistiques de spécialité sur l'axe syntagmatique et sur l'axe paradigmatique ; variation sur le plan lexical, sur le plan phrastique, sur le plan textuel et sur le plan discursif » (Desmet 2007 : 4). Dans la présente étude, nous nous intéresserons à la variation des termes sur l'axe syntagmatique. Les termes n'existent jamais dans le vide : dans les textes spécialisés, ils apparaissent toujours dans un contexte donné, plus ou moins figé. Dans la terminologie actuelle, de plus en plus d'études se focalisent sur la problématique du contexte, c'est-à-dire sur ce qui apparaît – et avec quelle régularité – à gauche et à droite du terme (voir notamment Alexandru et Gaudin 2006, ainsi que les autres articles du même volume). Notre étude portera sur les termes et leur entourage syntaxique dans des textes spécialisés très spécifiques, à savoir ceux traitant de l'histoire de l'art, et de l'histoire de la peinture plus particulièrement. Dans le domaine de l'histoire de l'art, nous distinguons trois types de textes : ceux écrits par des spécialistes et destinés aux spécialistes, ceux écrits par des spécialistes et destinés principalement aux étudiants du domaine (donc à visée didactique), et ceux écrits par des spécialistes et accessibles non seulement aux spécialistes, mais aussi au large public. Dans notre analyse, nous nous concentrerons sur ce troisième type de textes, dans lesquels les spécialistes (les historiens de l'art) – à travers l'histoire de la peinture – abordent la problématique de la couleur en essayant de décrire les nuances des couleurs utilisées par tel ou tel peintre à telle ou telle époque.

Notre étude se concentrera sur le contexte de deux termes de couleur choisis : *rouge* et *vert*. Notre objectif est de vérifier s'il existe des régularités et des répétitions dans ce qui peut accompagner ces termes dans les textes français spécialisés relevant de l'histoire de l'art : quels sont les verbes ou les adjectifs qui accompagnent le plus souvent ces termes de couleur ? Dans les textes étudiés, les vocables *rouge* et *vert* peuvent désigner soit une couleur-matière (ex. « Le rouge minium (oxyde de plomb) n'apparaît qu'à l'époque romaine »), soit une couleur perçue (ex. « (...) mais on voit également deux verts (un vert foncé et un vert un peu turquoise) (...) »), éventuellement une couleur spectrale. Les questions auxquelles nous chercherons à répondre dans cette communication sont les suivantes : y a-t-il des caractéristiques syntaxiques propres à chacun de ces usages et de ces significations ? Y a-t-il un éventail plus ou moins figé d'adjectifs ou de verbes qui accompagneraient ces deux couleurs ? Si on trouve une quelconque régularité, peut-on parler déjà de spécialisation de ces unités accompagnatrices ? Peut-on définir les affinités sémantiques qui sous-tendraient l'apparition d'une certaine combinatoire des termes dans le domaine en question (L'Homme 2004 : 111-114) ? Notre analyse des termes désignant les couleurs et de leur proche contexte visera à répondre à toutes ces questions en prenant surtout en compte ce caractère double du phénomène de la couleur, situé entre la matière et l'imagination.

### Références bibliographiques

- Alexandru, C., Gaudin F. (2006). Les contextes : à la source du terme ? In Blampain D., Thoiron Ph., Van Campenhoudt M. (éds.), *Mots, termes et contextes. Actes des septièmes Journées scientifiques du réseau de chercheurs Lexicologie Terminologie Traduction*. Paris : Éditions des archives contemporaines, 59-67.
- Desmet, I. (2007). Terminologie, culture et société. Éléments pour une théorie variationniste de la terminologie et des langues de spécialité. *Cahiers du Rifal*, 26, 3-13.
- L'Homme, M.-C. (2004). *La terminologie : principes et techniques*. Montréal : Les Presses de l'Université de Montréal.
- Mejri, S. (2018). La phraséologie : cotexte, contexte et contenus culturels. *Lublin Studies in Modern Languages and literature*, 42(4), 11-38.
- Pavel, S. (1993). La phraséologie en langues de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques. *Terminologies Nouvelles*, n° 10, 67-82.

## Vocabulaire de la chimie : l'analyse sémantique des phrasèmes

Polina Mikhel (ATILF-CNRS, Université de Lorraine)  
 Tomara Gotkova (ATILF-CNRS, Université de Lorraine)  
 Francesca Ingresso (LPCT-CNRS, Université de Lorraine)

Depuis 2018, nous travaillons sur la construction de modèles lexicographiques de termes de chimie générale et de chimie verte, modèles intégrés à la description du lexique de la langue générale dans le cadre des travaux menés au laboratoire ATILF (Analyse et traitement de la langue française) sur les grands réseaux lexicaux appelés *Systèmes Lexicaux*. L'objectif de notre projet est de créer une ressource trilingue sur le lexique de la chimie qui vise à devenir une ressource de référence pour les étudiants et les enseignants de la chimie, ainsi que pour les traducteurs français-anglais-russes.

Nos recherches reposent sur l'hypothèse que les terminologies scientifiques, puisqu'elles fonctionnent dans les textes en interaction avec le lexique de langue générale, possèdent une structure homomorphe avec celle du lexique général, avec lequel elles fusionnent au sein de la langue (L'Homme, Polguère, 2008). Notre approche théorique et descriptive est celle de la *Lexicologie Explicative et Combinatoire* (Mel'čuk *et al.*, 1995). Le travail lexicographique et terminographique est mené de façon comparative pour trois langues typologiquement distinctes : français, anglais et russe.

Dans le cadre de nos travaux, les termes chimiques reçoivent le même traitement que les unités de langue générale : on propose la description terminographique détaillée de leur sens, caractéristiques grammaticales, connexions lexicales sémantiques et phraséologiques ; les exemples d'utilisation des termes dans les textes scientifiques sont tirés des corpus textuels de chimie créés à l'ATILF dans le cadre de deux projets coordonnés<sup>1</sup>. L'attention se porte particulièrement sur l'interface entre langue générale et langue de spécialité, autour des unités lexicales générales – les non-termes<sup>2</sup> – et des termes de chimie qui établissent cette jonction, surtout au niveau de la polysémie de vocables (*substance, molécule, liaison, réagir, catalyser*, etc.).

Dans ce contexte, on s'intéresse aux phrasèmes, et tout particulièrement aux collocations. On entend par collocation un syntagme *AB* (ou *BA*) qui est tel que, pour le construire, le Locuteur sélectionne *A* – une *base de collocation* – librement d'après son sens 'A', alors qu'il sélectionne *B* – un *collocatif* – pour exprimer auprès de *A* un sens 's' en fonction de contraintes imposées par *A* (Polguère 2016).

Au cours de nos travaux, nous avons identifié un nombre de syntagmes qui, à première vue, semblaient être tous des collocations impliquant les termes et non-termes. Néanmoins, après l'analyse sémantique nous avons établi qu'ils appartiennent aux trois types de syntagmes suivants :

- 1) **syntagmes libres**, qui sont composées de deux termes distincts ayant une grande affinité statistique et qui sont donc décrits dans des articles lexicographiques séparés ; par exemple, *molécule neutre* : terme de chimie *molécule* 'groupement d'atomes liés...' + terme de physique *neutre* 'qui n'est chargé ni positivement ni négativement' ;
- 2) **collocations**, qui ne représentent pas un terme unique ; la base de collocation et le collocatif sont décrits dans les articles lexicographiques séparés, mais sont connectés à l'aide des fonctions lexicales (voir Mel'čuk *et al.*, 1995 et Mel'čuk, Polguère, 2020) ; par exemple, *liaison forte* : terme de chimie *liaison* 'forces qui assurent le lien entre les atomes...' + lexème de langue générale *fort(e)* 'tel que ça doit être' (fonction lexicale **Ver**) ;
- 3) **locutions**, y compris des **locutions faibles** (Polguère, 2015 ; Abreu, 2018), qui représentent un concept du domaine de chimie et qui sont décrits dans un article lexicographique unique ; par exemple, *poids atomique* 'rapport de la masse moyenne d'un atome au douzième de la masse d'un atome de carbone 12...'.

Dans notre exposé, nous allons présenter notre approche théorique et descriptive, donner des exemples de divers types de phrasèmes du lexique de chimie générale et de chimie verte, et expliquer comment on les représente dans nos modèles lexicographiques. Finalement, nous allons traiter l'application éventuelle de cette approche dans le contexte de l'enseignement de la chimie.

<sup>1</sup> Projet *FRONTERME*, financement d'une recherche doctorale par la région Grand Est, France ; projet *LEGCOD*, financement d'une recherche doctorale par IMPACT OLKi, Nancy, France.

<sup>2</sup> D'après L'Homme et Polguère, « les termes se distinguent des unités lexicales générales — les unités lexicales non marquées — uniquement par le fait que leur maîtrise linguistique est associée à la maîtrise d'un domaine de connaissance donné (scientifique, technique, etc.) » (L'Homme, Polguère, 2008).

## Références bibliographiques

- Abreu, S. P. (2018). Considerações sobre a natureza colocacional e locucional de sintagmas terminológicos. *ReDILLeT*, 1(2), en ligne.
- Gader, N., Lux-Pogodalla, V., Polguère, A. (2012). Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III)*, 109-125, 109-126. The COLING 2012 Organizing Committee.
- Ingrosso F., Polguère A. (2015). How Terms Meet in Small-World Lexical Networks : The Case of Chemistry Terminology. In Poibeau, T., Faber, P. (eds.), *Proceedings of the 11<sup>th</sup> International Conference on Terminology and Artificial Intelligence (TIA 2015)*, 167-171.
- Jasien, P.G. (2011). What Do You Mean That “Strong” Doesn’t Mean “Powerful”? *Journal of Chemical Education*, 88 : 9, 1247-1249.
- L’Homme, M.-C. (1998). Caractérisation des combinaisons lexicales spécialisées par rapport aux collocations de langue générale. In Fontenelle, T. et al. (eds.), *Actes du congrès « Proceedings Euralex ’98 »*, 513-522. Université de Liège.
- L’Homme, M.-C. (2003). Les combinaisons lexicales spécialisées (CLS). Description lexicographique et intégration aux banques de terminologie. In Grossmann, F., Tutin, A. (eds.), *Les collocations : analyse et traitement, Travaux et Recherches en Linguistique Appliquée*, 89-105. De Werelt.
- L’Homme, M.-C. (2007). Using Explanatory and Combinatorial Lexicology to Describe Terms. In Wanner, L. (ed.), *Selected Lexical and Grammatical Topics in the Meaning-Text Theory. In Honour of Igor Mel’čuk*, 163-198. John Benjamins.
- L’Homme, M.-C., Polguère, A. (2008). Mettre en bons termes les dictionnaires spécialisés et les dictionnaires de langue générale. In Maniez, F., Dury, P. (eds.), *Lexicologie et terminologie : histoire de mots. Hommage à Henri Béjoint*, 191-206. Travaux du CRTT.
- Mel’čuk, I., Clas, A., Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot.
- Mel’čuk, I., Polguère, A. (à paraître). Les fonctions lexicales dernier cri. In Marengo S. (ed.), *La Théorie Sens-Texte et ses applications. Lexicologie, lexicographie, terminologie, didactique des langues*, coll. « Dixit Grammatica ». L’Harmattan.
- Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27(4). Special Issue : *Dictionaries and the Digital Revolution : A Focus on Users and Lexical Databases*, 396-418.
- Polguère, A. (2015). Non-compositionnalité : ce sont toujours les locutions faibles qui trinquent. *Verbum*, II, XXXVII (2), 257-280.
- Polguère, A. (2016). *Lexicologie et sémantique lexicale. Notions fondamentales*. Troisième édition. Presses de l’Université de Montréal.

## Les moyens d’enrichissement du vocabulaire des coloronymes cosmétiques français et leurs équivalents russes et italiens

Galina Ovchinnikova (Université d’État pour la région de Moscou)  
Assya Ovchinnikova (Académie Internationale de Police, Toula)

Le but de cette intervention est de continuer les recherches dans le domaine du système terminologique de la mode, commencées depuis le siècle dernier par la thèse de R. Barthes (1967) et par la thèse de A. J. Greimas (1948). Ce sujet ne cesse d’intéresser les linguistes des pays différents jusqu’à présent (Mathe 2014, Cavaille-Fol 2020), dont les recherches de Michel Pastoureau (2013, 2014a, 2014b, 2019a, 2019b) sont les plus importantes. La palette de couleurs attire l’attention des linguistes non seulement dans le domaine vestimentaire, mais aussi comme une catégorie universelle (Cavaille-Fol 2020), car les couleurs façonnent notre perception. Certains linguistes remarquent qu’il existe aussi d’importantes divergences dans les systèmes des noms de couleur dans les langues et les cultures différentes et analysent les termes chromatiques du point de vue traductologique (Krylosova 2019).

Nous espérons dresser un « portrait » des multiples possibilités offertes par les systèmes dérivationnels français, italien et russe de fabriquer des unités lexicales chromatiques en cosmétologie et d’aider à mieux comprendre le processus de créativité lexicale. En nous basant sur l’analyse étymologique, morphémique, dérivationnelle, nous visons à soumettre aux discussions des terminographes une nouvelle classe lexico-grammaticale du champ sémantique « mode » qui concerne le vocabulaire des coloronymes dans le domaine cosmétique. On introduit pour la première fois le terme « un coloronyme cosmétique », en précisant ses différences par rapport au « nom de couleur », « terme de couleur », « terme chromatique » et d’autres. La particularité des coloronymes, à la différence des termes cités, est de présenter une couleur ou une nuance de couleur, obtenue par des moyens artificiels, chimiques à la suite du mélange des couleurs.

Nous avançons une nouvelle approche à l’étude typologique de ces coloronymes dans la sphère cosmétologique en trois langues : le français, l’italien et le russe. Grâce à l’analyse componentielle, on fait la classification des coloronymes cosmétiques selon les principes de l’enrichissement du vocabulaire coloronymique en cosmétologie dans les trois langues des familles différentes. L’étude du corpus des coloronymes (qui embrasse 350 termes français, 420 termes italiens et 320 termes russes des années 2017-2020, tirés des dictionnaires bilingues et trilingues, des données du Terminalf du Laboratoire de Loïc Depecker, des catalogues cosmétiques en ligne, des pubs télévisées) montre que l’enrichissement du vocabulaire coloronymique français et italien en cosmétique s’effectue en premier lieu au moyen de la conversion. Certains linguistes l’appellent la dérivation impropre ou la dérivation à suffixe zéro.

- le rouge à lèvres : fraise, cerise, chocolat, noisette, marron
- i rossetti: fragola, ciliegia, cioccolato, nocciola, marrone
- губная помада: клубника, шоколад, лесной орех, каштан, коралловое облако, элегантная фуксия, розовый персик, пряная карамель, сочный коралл, пламенный красный, вишневый бархат.



En ce qui concerne l'exemple *вишневый бархат* /velours cerise et d'autres, les résultats des recherches traductologiques de Světlana Krylosova nous ont servi de base de la formation des fichiers de la définition pour le futur dictionnaire trilingue.

C'est seulement les dernières années que la conversion participe activement à la formation des coloronymes cosmétiques russes, mais on observe la tendance d'ajouter une épithète ce qui provoque la formation des syntagmes lexicalisés.

La place prépondérante dans la formation des coloronymes cosmétiques russes appartient à la dérivation suffixale : *неб-ес-ный* (*bleu ciel*), *жемчуж-н-ый* (*perle*), *лазур-н-ый* (*azur*)

Les séries dérivationnelles sont typiques en français pour le modèle : adjectif + le suffixe -âtre : *vert – verdâtre*, en russe et en italien on recourt plutôt aux mots composés : *светло-зеленый, verde chiaro*.

Il y a un groupe structurel de coloronymes qui présente un grand intérêt dans les trois langues : *couleur oxide de fer – colore ossido di ferro – окись железа, bleu de Delft – blue di Delft – дельфт блю (emprunt), gris d'acier – grigio acciaio – серый стальной*.

La question se pose de savoir comment les nommer : lexies, mots composés, combinaisons syntaxiques. Nous proposons le terme « les syntagmes lexicalisés » pour ce groupe de coloronymes, car du point de vue du sens ils ont une signification qui se réalise en plusieurs formes.

Nous cherchons à répondre à la question comment les mots se sont constitués ou reconstitués et quelles règles sont susceptibles de rendre compte de la création permanente des nouveaux lexèmes et lexies, en démontrant que les coloronymes de toutes les langues en question font partie des systèmes créateurs de formation de nouveaux mots selon le principe : A de nouveaux besoins, de nouveaux mots.

Bien sûr, chaque époque demande son dictionnaire de coloronymes. L'analyse sémique nous a permis de faire la classification des coloronymes français, italiens et russes dans le domaine cosmétique en groupes suivants :

- 1) Les coloronymes qui contiennent le sème différentiel « toponyme » : *Bordeaux, fr., it., Бордо ; l'azur de Berlin – azzurro Berlino-берлинская лазурь chartreuse, violetto veneziano, giallo Napoli, rosso corsa*.
- 2) Les coloronymes qui contiennent le sème différentiel « floronyme » : *fuchsia, fr., it., -фуксия, rose-rosa-розовый, foglia di tè, magnolia*.
- 3) Les coloronymes qui contiennent le sème différentiel « zoonyme » : saumon (saumon, rose-orange vif) - *salmone -цвет лосося, canard / bleu canard (bleu-vert) – blu pavone – сизый павлиний, canari (jaune canari) – giallo canarino – канареечно жёлтый, grenouille (vert foncé) verde rana -лягушачий зелёный, caméléon (beige brunâtre) – цвет – beige camaleonte – цвет хамелеон, grigio topo, grigio souris*,
- 4) Les coloronymes qui contiennent le sème différentiel « produit alimentaire » : *café au lait – caffè latte – кофе с молоком, caramél – caramello-карамельный, rosa arancio, orange, aubergine – баклажан viola melanzana, crème – смета -кремовый, miel – miele – медовый*.
- 5) Les coloronymes qui contiennent le sème différentiel « noms des pierres et des minéraux » : *améthyste (violet clair) – ametista – аметистовый, émeraude – smeraldo – изумрудный, cobalt-cobalto-кобальтовый, platine-platino-платиновый, antracite-антрацит, charbon*.
- 6) Les coloronymes qui contiennent le sème différentiel « phénomène de nature » : *horizon (bleu-gris) – blu-grigio- серо-голубой, aurore (rose doré) – aurora-багряный, feu (rouge-feu) – rosso fuoco- огненно-красный*
- 7) Les coloronymes qui contiennent le sème différentiel « une personnalité » : *голубой/розовый Пикассо – bleu/blu, rose/rosa Picasso, Gainsboro, rosso Tiziano, verde Veronese*

Il est important de soumettre à l'attention des chercheurs une perspective d'étude des locutions dont un des composants est un coloronyme, par exemple : *un bonbon rose* a un sens péjoratif (une personne d'un comportement peu sérieux).

La présente communication s'inscrit aussi dans le courant des recherches phraséologiques qui se focalisent sur l'aspect lexicographique et contrastif des expressions figées. Nous essayons de préciser l'appareil notionnel phraséologique qui porte sur les idiomes, les unités phraséologiques, ainsi que de résoudre le problème d'équivalence des locutions qui est largement discuté, mais rarement analysé sur un plan purement phraséologique. Le corpus d'expressions figées avec un coloronyme est plus restreint avec 200 locutions en français, italien et russe, mais il est assez représentatif.

Dans le cadre de cette communication nous nous concentrons plutôt sur l'équivalence sémantique, fonctionnelle et référentielle, en traitant l'équivalence formelle d'une façon moins rigide, en nous limitant aux coloronymes, composants des locutions, les plus répandus : noir, gris, blanc, rouge, bleu, vert, jaune et leurs équivalents italiens et russes.

L'analyse du corpus illustratif des expressions figées chromatiques dans le plan sémique a montré que ces dernières, à la différence des coloronymes cosmétologiques, peuvent avoir des sèmes connotatifs avec une nuance positive, mais très souvent avec une nuance négative :

- *voir noir, le jour noir – видеть в черном цвете, черный день*.
- *essere nero – выглядеть чернее тучи*.
- *il libro nero – черный список*.
- *una giornata grigia – серый ден*.
- *vita grigia – монотонная жизнь*.
- *imore grigio – тоскливое или упадочническое настроени*.
- *ore grigie – тоскливое время, тоскливая пора*.

Les locutions italiennes préfèrent le coloronyme *brun*, où les Russes et les Français utilisent le coloronyme *noir* :

- *aver il bruno – носить траур*.
- *il bruno, il bel non togliè - хоть и черен, да мил*.

- *donna e luna, oggi è serena, domani è bruna* – женский нрав переменчив.
- *bruno al cuore* – мрачное настроение.

Les coloronymes *blanc, vert, rouge, jaune* et *bleu* forment plus de locutions avec des sèmes connotatifs positifs :

- *blanc comme la neige, bianco come il latte* – белее снега.
- *anni verdi* – зеленые годы, *verde vecchiezza* – счастливая старость, *giallo come un limone, jaune comme un citron* – желтый как лимон, *avoir du sang bleu, avere il sangue blu* – быть аристократом, голубых кровей.

Ces coloronymes forment aussi beaucoup de locutions avec les sèmes connotatifs chromatiques négatifs :

- *cousu de fil blanc* – шито белыми нитками, *restare bianco e antico* – остаться ни с чем, *diventare verde*.
- *di rabbia/ invidia* – позеленеть от злости / зависти (*farsi verde*).
- *c'est chou vert et vert chou* – что в лоб, что по лбу, *vedere rosso* – быть злым, *avoir un pied de rouge sur la figure* – раскраситься, *un vin bleu* – плохое вино, *une envie bleue* – чёрная зависть, *avoir une colère bleue* – сильно злиться, *un rire jaune* – неискренний смех, *giallo di paura* – сильно бояться.

Après avoir analysé notre corpus des locutions dans une perspective contrastive, il est possible d'admettre que les langues des familles différentes restent en général également plus proches sur le plan des coloronymes et des locutions avec le composant chromatique.

Le corpus du matériel illustratif des coloronymes cosmétiques, qui compte 1110 termes, prouve que les moyens d'enrichissement coloronymique du vocabulaire cosmétique français et italien sont la conversion : le substantif acquiert les fonctions adjectivales et la formation des syntagmes lexicalisés. Par contre, le russe reste fidèle à la dérivation nominale, mais la conversion commence à prendre son terrain dans la formation du système terminologique des cosmétoponymes de couleur en russe.

Les résultats des recherches obtenus permettent de continuer à élaborer les principes du fonctionnement de ces termes dans les textes littéraires des siècles différents et à voir l'évolution de la formation de ces termes. Le matériel factitif analysé est prêt à la dictionnarisation des coloronymes des langues des familles langagières différentes.

Les approches proposées peuvent servir d'outils de recherche dans d'autres langues et non seulement en cosmétologie, mais dans d'autres systèmes terminologiques.

## Références bibliographiques

- Barthes, R. (1967), *Système de la mode*. Points.
- Cavaille-Fol, T. (2020) Couleurs : elles sont gravées dans notre cerveau, <https://www.science-et-vie.com/cerveau-et-intelligence/couleurs-elles-sont-gravees-dans-notre-cerveau/53620>
- Greimas, A.-J. (1948), *La mode en 1830. Langage et société*. PUF.
- Mathe, A. (2014). Sémiotique du vêtement, aujourd'hui. *Actes sémiotiques*, 117/2014, en ligne, <https://www.unilim.fr/actes-semiotiques/4964>
- Fracassi, P., Marotolli, M.S. (2006). *Dizionario di dermocosmesi. 1250 termini inglese-italiano e italiano-inglese*, Tecniche Nuove.
- Krylova, S. (2009). Sémantique des termes chromatiques et problèmes de traduction. In *La Traduction : philosophie, linguistique et didactique. Actes du colloque international (1-3 avril 2009, Lille)*, Université Charles-de-Gaulle-Lille 3, 229-232.
- Krylova, S. (2010), « Lampa pod vichněvym platkom : o perevode odnogo ». *Histoire d'une couleur*, Seuil, 286-290.
- Krylova, S. (2010), « Ces yeux couleur vinaigre. Adjectifs chromatiques dans les dictionnaires bilingues : dialogue des cultures ? », *Le lexique entre identités culturelles et genres textuels. Colloque international (11-13 novembre 2010)*, Università degli Studi di Verona).
- Ovtchinnikova, G. (2015), *Konversiya kak sposob slovoizvodstva leksicheskogo plasta « moda » v sovremennom frantsuzskom yazyke*. Toula.
- Pastoureau, M. (2014a), *Bleu. Histoire d'une couleur*. Seuil.
- Pastoureau, M. (2014b), *Noir. Histoire d'une couleur*. Seuil.
- Pastoureau, M. (2013), *Vert. Histoire d'une couleur*. Seuil.
- Pastoureau, M. (2019a), *Rouge. Histoire d'une couleur*. Seuil.
- Pastoureau, M. (2019b), *Jaune. Histoire d'une couleur*. Seuil.
- TERMINALF. Ressources terminologiques en langue française. <https://terminalf.scicog.fr/>

## Méthode contrastive et multifactorielle de repérage automatisé des expressions multilexémiques lexicalisées

Éric Poirier (Université du Québec à Trois Rivières)

Notre proposition de communication s'inscrit dans l'axe numéro 2 sur le traitement automatique des langues. Nous expliquons comment exploiter les très nombreux corpus parallèles de traduction afin de repérer des expressions multilexémiques lexicalisées (de la langue générale, donc phraséologiques, et aussi de la langue spécialisée, donc terminologiques) grâce à une méthode automatisée d'analyse multifactorielle des écarts de différents types de densités lexicales entre les paires de segments source et cible. La méthode décrite vise la découverte d'expressions multilexémiques lexicalisées, au sens de Constant, M., Eryigit, G., Monti, J., et autres (2017), par opposition à l'identification d'expressions déjà connues. L'analyse contrastive des paires de segments bilingues d'un corpus parallèle

peut s'associer à d'autres techniques de repérage statistique des unités phraséologiques fondées sur la fréquence des cooccurrents lexicaux dans des corpus unilingues, comme celle proposée par Colson (2017). Dans cet appariement, la méthode peut renforcer le statut multilexémique d'expressions découvertes avec d'autres techniques.

Le dispositif d'analyse empirique des paires de segments que nous décrivons ci-après s'inscrit dans le prolongement de la méthode de comparaison des volumes d'information que nous avons proposée récemment (Poirier 2019) dans l'exploitation de textes traduits de l'anglais au français. Un des avantages de notre méthode est le fait qu'elle s'applique à la plupart sinon à la totalité des langues qui font la distinction entre les sens et les mots lexicaux (une classe ouverte), et les sens et les mots grammaticaux (une classe fermée), lesquels ne contribuent pas de la même façon au sens des énoncés et sont par conséquent traduits d'une manière très différente (Polguère 2003 : 107-108, Poirier 2003).

L'analyse des écarts de densité lexicale à partir de corpus bilingues parallèles de traduction nous semble tout particulièrement justifiée pour repérer les unités phraséologiques opaques, celles dont le sens n'est pas déductible de leurs éléments constitutifs et qui ont la particularité d'être traduites par une périphrase sémantique fonctionnelle (non phraséologique) qui est susceptible de compter un nombre différent (réduit ou accru) de sens ou de mots lexicaux. Par contre, elle est probablement moins indiquée pour repérer d'autres catégories d'unités phraséologiques comme les collocations qui se manifestent rarement par des écarts de densité lexicale.

Même si elles sont vraisemblablement plus nombreuses que les autres unités de la langue (Mel'Čuk 2013), les unités phraséologiques représentent des exceptions ou des irrégularités syntaxiques de par leur caractère non libre ou figé. En tant qu'expressions (multilexémiques), elles deviennent ainsi des objets du lexique, par leur figement et les contraintes syntaxiques qui s'appliquent à elles (Gross, 1996). La méthode multifactorielle et contrastive que nous proposons s'appuie sur le repérage de deux conséquences courantes (mais pas systématiques) de leur lexicalisation (l'expansion ou la réduction lexicale) qui sont constatées en traduction par l'écart des densités lexicales relatives mesurées automatiquement entre les segments source et cible d'un corpus parallèle de traduction. Il serait possible d'étendre cette analyse bilingue à des corpus parallèles multilingues de manière à accroître les possibilités d'occurrence des écarts de densité lexicale relative (expansions et réductions lexicales), et de repérage des unités phraséologiques, et à corroborer plus fortement la nature idiosyncrasique de l'unité phraséologique mise en évidence de manière contrastive.

Notre méthode repose sur le calcul automatique de deux catégories de ratios linguistiques relatifs pour chaque segment source et cible de la paire de segments de chaque texte versé dans un corpus de traduction parallèle. Le volume d'information relatif des segments source et cible est calculé en normalisant le nombre de mots lexicaux dans les segments en fonction du segment type du texte et en tenant compte aussi du nombre de mots lexicaux qui se trouve dans l'autre segment de la paire analysée. De même, le ratio de foisonnement relatif des segments source et cible est calculé en normalisant le nombre de mots dans les segments en fonction du segment type du texte et en tenant compte également du nombre de mots totaux qui se trouve dans l'autre segment de la paire analysée. Une fois ces quatre données compilées pour chaque segment, on calcule la distance euclidienne entre les segments A et B grâce au volume d'information relatif de chaque segment et au ratio de foisonnement relatif de chaque segment. À titre d'exemple choisi parmi les segments repérés avec cette méthode d'analyse du discours d'Obama du 22 mars 2016 prononcé à Cuba en anglais et traduit en espagnol, on trouve la paire de segments numéro 138 ci-dessous qui affiche une distance euclidienne des densités lexicales relatives de 3,74 (sur une échelle comprise entre 0,04 et 7,34 pour l'ensemble des paires de segments du discours) :

*Not everybody agrees with the American people on this. = No todo el mundo está de acuerdo con el pueblo estadounidense sobre esto.*

L'analyse manuelle des sens lexicalisés dans la paire de segments met en évidence dans le segment en espagnol la présence de deux unités phraséologiques, à savoir *todo el mundo* ainsi que *estar de acuerdo* qui expliquent l'hétéromorphisme de la traduction avec le segment source. Les deux expressions ont des correspondants multilexémiques en français, soit « tout le monde », qui contient un adjectif et un nom et qui est considéré comme une locution nominale dans le Multidictionnaire, et « est d'accord » qui contient un verbe, auxiliaire, et une chaîne de caractères un peu particulière, « d'accord », qui est considérée comme une locution adverbiale formée d'une préposition et d'un nom). À titre de comparaison avec un autre segment du même texte qui ne répond pas favorablement aux critères de notre analyse, et qui est ainsi considéré comme étant isomorphe, on trouve le segment no 21 ci-dessous qui affiche une très faible distance euclidienne de 0,04 (malgré la longueur des segments qui totalisent respectivement 19 et 18 mots lexicaux et grammaticaux) et qui ne contient vraisemblablement pas d'unité multilexémiques :

*Those waters also carried generations of Cuban revolutionaries to the United States, where they built support for their cause. = Esas aguas también transportaron a generaciones de revolucionarios cubanos hasta Estados Unidos, donde consiguieron apoyo para su causa.*

Notre communication vise à présenter les résultats du traitement de textes spécialisés traduits du français à l'anglais et de l'anglais au français, et possiblement d'autres langues comme l'espagnol ou l'italien, dans le but de repérer de façon discriminante des unités multilexémiques du français. D'autres moyens complémentaires ou plus spécialisés pourront être mis en œuvre dans l'amélioration de notre dispositif d'analyse, comme l'utilisation des ratios de fréquence dans les corpus unilingues qui sert à confirmer la nature phraséologique ou spécialisée (terminologique) des expressions multilexémiques, la prise en compte d'autres traductions des mêmes segments dans des langues différentes grâce à des corpus parallèles multilingues, la prise en compte de la traduction automatique des segments source et cibles, par comparaison avec la formulation de départ ou la ou les traductions officielles, la comparaison des résultats des mêmes analyses menées dans des corpus parallèles en langue de spécialité avec les résultats obtenus dans des corpus parallèles

de langue générale, ce qui pourrait fonder une partie de la distinction utile entre unités multilexémiques terminologiques et unités multilexémiques phraséologiques. Enfin, l'exploration des propriétés syntaxico-sémantiques de certaines constructions telles que les contraintes de sélection sémantiques, les propriétés grammaticales et les patrons syntaxiques, comme celles qui ont été décrites dans Poirier (2015), pourraient aussi s'intégrer à la méthode d'analyse contrastive multifactorielle que nous proposons dans l'évaluation des écarts entre paires de segments parallèles, ce qui permettrait de repérer à partir des textes traduits des idiosyncrasies expressives et des unités phraséologiques typées et catégorisables en fonction de l'acception de leur principal élément lexical constitutif, de leur acception globale et de leur nature prédicative (verbale, nominale, adjectivale ou autre), ou non prédicative.

## Références bibliographiques

- Colson, J.-P. (2017). The IdiomSearch Experiment : Extracting Phraseology from a Probabilistic Network of Constructions. In Mitkov R. (dir.) *Computational and Corpus-Based Phraseology. EUROPHRAS 2017*. « Lecture Notes in Computer Science », vol 10596. Springer : Cham.
- Constant, M., Eryiğit, G., Monti, J., et autres (2017). Multiword expression processing : A survey. *Computational Linguistics*, 43(4), 837-892.
- Gross, G. (1996). *Les expressions figées en français : noms composés et autres locutions*. Paris : Éditions Ophrys.
- Mel'čuk, I. (2013). Tout ce que nous voulions savoir sur les phrasèmes, mais.... *Cahiers de lexicologie*, 102(1), 129-149.
- Obama, B. (2016). *Remarks by President Obama to the People of Cuba*, The White House President Barack Obama, <https://obamawhitehouse.archives.gov/the-press-office/2016/03/22/remarks-president-obama-people-cuba>.
- Obama, B. (2016). *Discurso del Presidente Obama al Pueblo Cubano*, Univision, <https://www.univision.com/noticias/obama-en-cuba/discurso-del-presidente-barack-obama-en-la-habana-cuba>
- Poirier, É. (2003). Conséquences didactiques et théoriques du caractère conventionnel et arbitraire de la traduction des unités phraséologiques. *Meta: journal des traducteurs/Meta: Translators' Journal*, 48(3), 402-410.
- Poirier, É. (2015). The interpretation of business in specialized expressions and compound terms for translation purpose. *InTRAlinea: Online Translation Journal*.
- Poirier, É. A. (2019). Repérage des décalages informationnels de traduction au moyen du criblage automatique des segments hétéromorphes d'un corpus parallèle, *TTR Traduction, Terminologie, Rédaction*, 32(1), 279-308.
- Polguère, A. (2003). *Lexicologie et sémantique lexicale, notions fondamentales*. Montréal : Les Presses de l'Université de Montréal.

## Présence et traitement des phrasèmes dans les vocabulaires juridiques français

Chiara Preite (Università di Modena e Reggio Emilia)  
Daniela Dincă (Universitatea din Craiova)

Considérées comme des « phrasèmes » en vertu de leur caractère non compositionnel (Mel'čuk 1995, 1996, 1998, 2003), les collocations ont fait l'objet de nombreuses définitions et typologies dans la littérature de spécialité, ce qui témoigne de leur complexité et de la multitude des pistes d'analyse dans ce domaine.

Dans la présente communication, nous nous proposons d'apporter quelques éléments de réflexion sur le traitement des phrasèmes de type *collocation verbale* dans les vocabulaires juridiques français. La définition que nous adoptons est celle de Larivière, synthétisée de la manière suivante : « La collocation est une combinaison semi-figée, récurrente et arbitraire, d'unités lexicales qui ont développé une affinité ou relation privilégiée consacrée par l'usage, qui ont une combinabilité restreinte et un statut distinct, l'une étant sémantiquement autonome et l'autre n'acquérant son sens qu'au sein de la combinaison, et dont la somme n'égale que partiellement le sens individuel de chacune » (Larivière 1998 : 177). En plus, nous privilégions également les deux types de collocations verbales (Heid 1994 ; L'Homme et Bertrand 2000 ; Dinca et Preite 2019) : (1) *collocations lexicales*, qui partent de la base nominale pour préciser la liste des collocatifs verbaux et (2) *collocations conceptuelles*, qui prennent comme point de départ le verbe pour identifier ce que Lerat (2008) appelle « les classes d'objets ».

Le corpus de collocations verbales a été formé autour de deux classes d'objet : <PROCEDURE> (*action, poursuite, procès, recours*) et <ACTE JURIDIQUE> (*requête, réclamation, réquisitoire, plainte*) que nous avons inventoriées dans trois ouvrages lexicographiques français : le *Vocabulaire du juriste débutant* de Pierre Lerat (2017, 2<sup>e</sup> éd.), *Le Dictionnaire phraséologique plurilingue du droit* (2017) de Pierre Lerat et *Le Dictionnaire des cooccurrences* rédigé par Jacques Beauchesne.

À partir de ce corpus, nous nous proposons d'atteindre les objectifs suivants, qui se superposent également aux étapes de notre analyse :

1. inventorier par comparaison les collocations verbales recensées dans les trois vocabulaires français ;
2. mettre en évidence les informations lexicographiques apportées par les trois ouvrages pour la définition des collocations verbales (les schémas d'arguments et les classes d'objet) ;
3. souligner l'utilité d'un futur ouvrage lexicographique dédié exclusivement aux collocations verbales aussi bien pour le décodage que pour l'encodage dans une langue étrangère.

En fin de compte, nous voulons souligner l'idée que l'intérêt pour l'analyse lexicographique des phrasèmes se justifie pleinement par son prolongement dans des domaines connexes comme la traduction, le traitement automatique des langues ou la pédagogie des langues spécialisées, à savoir la langue du droit.

## Références bibliographiques

- Beauchesne, J. (2001). *Dictionnaire des cooccurrences*. Montréal : Guérin.
- Cornu, G. (1990). *Linguistique juridique*. Paris : Montchrestien.
- Cornu, G. (2016). *Vocabulaire juridique*. Paris : P.U.F.
- Darbelnet Jean (1979). Réflexions sur le discours juridique », *Meta*, 24(1), 9-17.
- Dechamps C. (2017). Pour un dictionnaire d'apprentissage des collocations en français et en portugais. *Synergies Brésil*, 12, 103-118.
- Dincă D., Preite C. (2019). Les collocations verbales dans le discours juridique : de la terminologie à la phraséologie. *Phrasis. Studi fraseologici e paremiologici*, 3, 136-147.
- Giráldez Ceballos-Escalera, J. (2010). La combinatoire collocationnelle dans le discours juridique : élément indispensable d'aide à la traduction. Communication au colloque *La traduction juridique - points de vue didactiques et linguistiques*, Lyon 3. [[http://www.ininterm.net/public/langues%20de%20sp%C3%A9cialit%C3%A9/colloque/Joaqu\\_n\\_Gir\\_ldez.pdf](http://www.ininterm.net/public/langues%20de%20sp%C3%A9cialit%C3%A9/colloque/Joaqu_n_Gir_ldez.pdf)].
- González Rey, M. I. (2002). *La phraséologie du français*. Toulouse : Presses Universitaires du Mirail.
- Guillien, R. et al. (éds.) (2016). *Lexique des termes juridiques*, Paris : Dalloz.
- Heid, U. (1994). On ways words work together - topics in lexical combinatorics. In *EURALEX094 Proceedings*. Vrije Universiteit Amsterdam, 227-257.
- Larivière, L. (1998). Valeur sémantique du verbe dans les collocations verbales spécialisées. *TTR : traduction, terminologie, rédaction*, 11(1), 173-197.
- L'Homme, M.-C. et Bertrand C. (2000). Specialized lexical combination s: should they be described as collocations or in terms of selectional restrictions?. *Proceedings Euralex'00. Ninth Euralex International Congress*. Stuttgart, 497-506.
- Lerat, P. (2008). Restrictions paradigmatiques et traduction de schémas d'arguments, *Meta*, 53(2), 434-442.
- Lerat, P. (2017). *Le Dictionnaire phraséologique plurilingue du droit*. Paris : La Maison du Dictionnaire.
- Lerat, P. (2017). *Vocabulaire du juriste débutant : décrypter le langage juridique*. Paris : Ellipses (2<sup>e</sup> éd.)
- Mel'čuk, I. (1996). Lexical Functions : A Tool for the Description of Lexical Relations in a Lexicon. In Wanner, L. (ed.). *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam : Benjamins Academic Publishers, 37-102.
- Mel'čuk, I. (1998). Collocations and lexical functions. In Cowie, A. P. (éd). *Phraseology. Theory, Analysis and Applications*. Oxford : Clarendon Press, 23-53.
- Mel'čuk, I. (2003). Collocations dans le dictionnaire. in Th. Szende (éd.), *Les écarts culturels dans les Dictionnaires bilingues*, Paris : Honoré Champion, 19-64.
- Mel'čuk, I. et al. (1995). *Introduction à la lexicologie explicative et combinatoire*. Bruxelles : Duculot.
- Perilli, F. (2009). Combinaisons lexicales et jurilexicographie. Le traitement des collocations terminologiques dans un dictionnaire de spécialité. In Chessa, F. et al. (éds.), *Les dictionnaires de spécialité. Une ouverture sur les mondes. Actes des Journées italiennes des dictionnaires. Troisièmes Journées, Cagliari, 3-4 octobre 2008*, Fasano-Paris : Schena-Alain Baudry & C<sup>ie</sup>, 227-243.
- Preite, C. (2017). La lexicographie juridique française comme véhicule de connaissance juridique et de compétence (juri)linguistique et communicative. *Travaux de linguistique*, 75(2), 101-123.
- Scurtu, G. (2008). Traduire le vocabulaire juridique français en roumain. *Meta*, 53(4), 884-898.

## Phénomènes phraséologiques dans le français juridique

Karina Rekosh (Institut de Moscou des Relations Internationales)

Le titre proposé reflète une approche multidisciplinaire : la phraséologie qui relève de la technique linguistique et le savoir juridique verbalisé par les termes conformes aux classifications des juristes. Le langage de spécialité se distingue de la langue générale : il est ordonné par le savoir qui conditionne sa terminologie, ses constances et immutabilité et s'oppose au langage courant, « dispersé » (Foucault 1969 : 100).

D'après les savants russes, la phraséologie distingue des groupements de mots libres et non libres (Teliya 1972 : 457), mais les dictionnaires juridiques ne le montrent pas. Nous les avons analysés d'après le dictionnaire *Lexique de droit commercial* (Bouilly 1992) et d'après le dictionnaire *Vocabulaire juridique* (Cornu 2009). Les dictionnaires juridiques contiennent les termes des textes (écrits) du langage linéaire qui verbalise le droit, son système conceptuel. La fonction communicative du langage juridique en état de dispersion n'est pas prise en compte.

Le plan de la recherche est le suivant : 1. Phraséologie comme théorie de la technique combinatoire de la langue qui a pour objet les groupements de mots non libres. 2. Mécanismes principaux de la langue (ordonnement et dispersion) et la phraséologie. 3. Méthode de décomposition des groupements de mots pour distinguer les collocations et locutions des groupements de mots libres (Tableau en annexe).

Les dictionnaires terminologiques résolvent (en partie) les problèmes de classifications lexicologiques. La terminologie s'étant détachée de la linguistique, les lexicologues ne sauraient, à notre avis, évaluer les professionnels en ce qui concerne les classifications des termes. Mais les juristes ne tiennent pas toujours compte des caractéristiques linguistiques des

termes, à l'exception de la « Linguistique juridique » de Cornu qui traite de la formation des mots composés (Cornu 2000 : 115) et des rapports entre les mots (Cornu 2000 : 174).

Il existe quelques critères d'étude des collocations basés : sur les transferts de sens si un composant est polysémantique (Teliya 1972 : 474) ; sur la base de distinction de signifié/signifiant pour tenir compte du nombre des référents ; sur les notions de langue/langage, ordonnancement/dispersion du langage, langage commun/ terminologique.

Les phénomènes phraséologiques n'existent pas qu'au niveau des syntagmes mais aussi au niveau du discours. Des adages du droit coutumier ont parfois un sens idiomatique qui ne découle pas du sens des mots : « Le tablier de la femme oblige le mari » (Roland et al. 1999 : 861). Les collocations sont soumises aux phénomènes d'univerbalisation, contraction des mots séparés en un seul sous les effets de conceptualisation et de sédimentation : ubiquité de la loi (cf. Ubi societas, ibi ius), mainmise, mainlevée. Les exemples cités montrent que les phrasèmes contiennent plusieurs mots mais le terme correspond à un référent et symbolise un seul concept qui guide les juristes, « faiseurs de systèmes » (cf. Bergel 2001). C'est le fonctionnement des mécanismes de sédimentation et de dispersion dans la langue qui la transforment en langage ordonné ou dispersé.

Notre méthode est destinée à vérifier, au niveau synchronique du langage ordonné, la divisibilité/indivisibilité des groupements avec ou sans changement de sens comme critères du figement ou de la liberté des bases ou collocatifs. C'est pratiquement l'étude des liens syntagmatiques où nous déterminons les paramètres du combinatoire, les liens paradigmatiques et de dérivation ne sont pas considérés. Le problème des collocations consiste dans la difficulté de les distinguer des groupements de mots libres et des locutions, surtout quand la composition est la même (cf. *dépôt de bilan/dépôt de garantie*). Nous proposons donc la décomposition des collocations en éléments pour voir le changement du sens (les parties du terme *dépôt de bilan* changent de sens, le sens des éléments de *dépôt de garantie* reste inchangé). Dans les dictionnaires phraséologiques, les idiomes et les collocations ne sont pas distingués. Les concepts linguistiques reflètent les relations verticales paradigmatiques, la sémantique de la langue générale (dispersée), ils font partie de la « pellicule fine de la civilisation » (Stepanov 2007) contrairement aux concepts juridiques faisant partie, avec les catégories, des constructions juridiques. « En matière juridique ils ne peuvent être déterminés qu'en fonction du droit objectif et des réalités sociales » (Bergel 2001 : 107).

Le tableau annexé montre que les structures des groupements de mots libres et non libres sont identiques et que la base de la collocation a, en règle générale, un sens conceptuel ou générique. Les collocatifs séparés de leurs bases peuvent aussi se transformer en concepts et figurer dans les dictionnaires isolément : *répétition de l'indu*. L'impossibilité de détacher la base du collocatif sans changement de sens témoigne du figement idiomatique. À l'instar des collocations, les collocatifs peuvent être composés de plusieurs mots et avoir différents degrés de figement, c'est-à-dire être constitués en idiomes ou collocations, à leur tour.

## Références bibliographiques

- Bergel, J.-L. (2001). *Méthodologie juridique*. Paris, P.U.F.  
Bouilly, M. (1992). *Lexique de droit commercial*. Paris : Hachette.  
Cornu, G. (2000). *Linguistique juridique*. Paris : Montchrestien.  
Cornu, G. (2009). *Vocabulaire juridique*. Paris, P.U.F.  
Foucault, M. (1969). *L'archéologie du savoir*. Paris : Gallimard.  
Roland, H. et al. (1999). *Adages du droit français*. Paris : Ed. Litec.  
Yu.S. Stepanov. (2007). Kontsepty. Tonkaya plenka tsivilizatsii. – Yazyki slavyanskikh kultur.  
Телия. Фразеология, (1972). Общее языкознание. Внутренняя структура языка. Изд. «Наука».  
Москва 1972, 565 с. V.N.Teliya. Frazеologiya. s. 456-515. Obshcheye yazykoznanie. Vnutrennyaya struktura yazyka. Izd. «Nauka».  
Moskva.  
Ю.С.Степанов. (2007). Концепты. Тонкая пленка цивилизации. – Языки славянских культур.

Paramètres des groupements de mots en fonction de\* leurs structures et de sens.

Groupements de mots		Divisibilité en éléments					
		Indivisibilité en éléments	Sens constant	Sens changé		Transferts de sens	
				entièrement	En partie	matériel	abstrait
<b>Libres</b>	Centres de gestion agréés Participation aux frais de transport Dépôt de garantie Action illicite sur les prix Tacite reconduction		+			Vente à la chandelle Départ pour Paris Abandon de poste	Départ à la retraite Abandon de créance
<b>Non libres</b>							
Idiomes	Pas de porte Cahier des charges Force majeure Fonds de commerce	+		+			
<b>Collocations :</b> Bases	Compte-rendu (de gestion) Autorité (de la chose jugée)				+		
Collocatifs	SSP (actes) sous seing privé (Salle) des pas perdus	+			+		
Structures :	Reçu pour solde de tout compte					+	
N+Adj Praepositio	Blanc-seing (donner) Nue-propriété Tierce opposition Convention collective Procès-verbal	+		+	+		
Postpositio	Motifs graves et légitimes Cause réelle et sérieuse Avoir fiscal Action publique Position dominante (abus de)		+			+	
N+N	Conseiller prud'hommes Chèque-resaurant Clause Molière Saisie-arrêt		+			+	
N+prép							
A	Action à droit de vote double Contrat à durée déterminée, indéterminée Travail à temps partiel, à plein temps				+		
De	Action d'apport Appel d'offres Dépôt de bilan Obligation de moyens Obligation de résultat			+	+	Jeton de présence	
En	Action en justice Action en répétition de l'indu Actions en numéraire				+		
Sur	Impôt sur les sociétés		+				
Par	Faute par commission Faute par omission		+				
Pour	Licenciement pour motif économique Intérêt pour agir Stipulation pour autrui		+				
Verbe + N Participe + N N + participe	Mettre en demeure Prête-nom Ayant cause Compte-rendu de gestion				+		

## Phrasèmes et collocations en langue de spécialité – le cas de la terminologie musicale ancienne

Delphine-Anne Rousseau (CRTT, Université Lumière Lyon 2)

Dans cette communication, nous aborderons la question des phrasèmes terminologiques, en les distinguant d'avec les collocations technolinguistiques, dans le contexte général de la terminologie en tant que discipline distincte de la lexicologie. Nous nous intéresserons à deux principaux aspects : les distinctions théoriques entre *phrasème* et *collocation* dans le contexte de la terminologie, ainsi que l'importance de la prise en compte des phrasèmes lors de l'extraction terminologique et de la conception d'ouvrages terminologiques.

### 1. Phrasème terminologique et collocation technolinguistique : considérations théoriques

Bien que des points de vue très divers aient été exprimés au sujet des collocations, phraséologismes et phrasèmes en langue générale, la frontière entre les notions de *collocation* et de *phrasème* en terminologie mérite encore d'être éclaircie. À notre avis, la principale différence entre ces deux notions réside dans le fait que la première ne renvoie pas, dans son intégralité, à un concept précis, contrairement à la seconde. Comme nous l'expliquerons, notre conception de *collocation technolinguistique* se rapproche de celle de Silva et al. (2004). Par contre, nous entendons par *phrasème* toute unité lexicale complexe dont au moins une partie est un terme mono- ou multisyntaxique, et dont le tout renvoie à un concept précis au sein du domaine. En ce sens, le phrasème constitue un terme à part entière. En ce sens, le *phrasème* correspondrait à la « [solidification de] la relation de dépendance entre les éléments de la construction » évoquée par Silva et al., et qui transforme selon elles une « *collocation terminologique* » en unité terminologique, bien que nous n'ayons pas été en mesure de vérifier diachroniquement la solidification en question pour les différents phrasèmes que nous avons relevés dans nos travaux. Le fait que nous soyons nous-même spécialiste du domaine de la musique ancienne a probablement beaucoup joué, entre autres choses, dans l'éclaircissement, ou plutôt dans la redéfinition des frontières entre *phrasème terminologique* et *collocation technolinguistique*. En d'autres termes, nos travaux nous ont permis d'aller au-delà de la simple compréhension théorique de la distinction entre ces deux concepts, en suscitant ce que nous appellerions une « compréhension tangible » de cette distinction.

### 2. L'importance de la prise en compte des phrasèmes et de toutes les catégories lexicales dans le travail terminologique

La terminologie est le reflet d'un domaine dans toute sa complexité. Toutefois, elle constitue aussi la porte d'entrée vers la complexité même de ce domaine. D'où l'intérêt de prendre en considération non seulement les phrasèmes, mais tous les termes d'un technolinguiste. C'est par l'élargissement de la collecte terminologique à toutes les catégories lexicales<sup>1</sup> porteuses de concepts (ce qui implique la fin de l'hégémonie substantivale dans le repérage des termes) que l'on peut non seulement brosser le véritable portrait d'une terminologie, mais également illustrer la variété qui existe effectivement dans la distribution des termes par catégorie lexicale entre différentes terminologies et comparer, au sein d'un même domaine, cette distribution entre différentes langues<sup>2</sup> (Rousseau 2019). Or, malgré tous les avantages que présente la prise en compte de l'ensemble des catégories lexicales dans l'étude des terminologies, nous avons constaté qu'elle est loin d'être systématique : ainsi, une grande partie des ouvrages terminologiques consultés au cours de nos recherches n'indiquent pas à quelle catégorie lexicale appartiennent les termes recensés ou définis, et seule la connaissance de la langue, ou dans certains cas la lecture de la définition, permet de distinguer dans ces ouvrages les noms des verbes, des adjectifs et des adverbes (voire les syntagmes des phrasèmes). Par ailleurs, certains des ouvrages multilingues qui précisent les catégories lexicales des termes ne les indiquent pas pour toutes les langues de l'ouvrage.

À l'évidence même, la prise en compte des différentes catégories lexicales s'avérerait hautement utile pour les chercheurs et pour la connaissance des domaines et des technolinguistes, en particulier pour les spécialistes qui ont à communiquer ou à comprendre un discours spécialisé dans une langue étrangère : il est en effet difficile de concevoir comment un expert d'un domaine (ou un expert en devenir) pourrait acquérir et maîtriser son technolinguiste en langue étrangère sans connaître, d'une part, la catégorie lexicale des termes et, d'autre part, les différentes collocations technolinguistiques. Mais cette prise en compte de l'ensemble des termes a une portée plus grande encore, car elle rend disponible les données permettant de calculer la distribution des termes par catégorie lexicale au sein des différents technolinguistes. Ces ratios permettent à leur tour de pousser plus loin l'étude et la compréhension des différents domaines, et ce d'au moins trois manières :

1) tout d'abord en rendant possible la comparaison entre différents domaines au sein d'une même langue : comme il a déjà été observé<sup>3</sup>, la nature d'un domaine ou d'une discipline se reflète souvent dans la distribution des termes par catégories lexicales, et on trouve par exemple souvent une représentation relativement élevée de verbes (et, surtout, de verbes d'action) dans les domaines ou disciplines où les actions occupent une place prépondérante ;

<sup>1</sup> C'est-à-dire celles qui correspondent à ce qu'on nomme les « mots pleins » ou « mots lexicaux » (noms, verbes, adjectifs et adverbes), mais également aux syntagmes nominaux, verbaux, adjectivaux et adverbiaux.

<sup>2</sup> Comme l'ont fait, notamment, les auteurs du *Handbook of business communication: linguistic approaches* (Mautner et al. 2017).

<sup>3</sup> Notamment dans Mautner et al. 2017 : 410-411.



2) ensuite, en permettant d'étudier l'évolution diachronique de ces ratios au sein d'un même domaine : cela permet de répondre à des questions comme « la nature d'un domaine varie-t-elle dans le temps ? » ;

3) enfin, ils offrent un nouveau prisme pour étudier les différences interlinguistiques pour un même technolecte : que dit la différence de distribution par catégorie lexicale entre deux langues au sein d'un même domaine ?

Ainsi, comme on peut le voir, la distribution des termes par catégorie lexicale au sein d'un technolecte donné traduit ou reflète le plus fidèlement la réalité du domaine auquel il appartient.

### **3. Phrasèmes terminologiques et collocations technolectes : quelques cas identifiés au cours de nos travaux**

Au cours de nos travaux sur l'étude de la terminologie musicale en usage en France et en Angleterre à la seconde moitié du XVII<sup>e</sup> siècle, fondés sur le dépouillement de plus d'une cinquantaine de traités et méthodes musicaux et de préfaces d'œuvres musicales de l'époque et des pays concernés, nous avons choisi non seulement d'inclure les phrasèmes relevés dans notre base de données, mais de relever également les termes appartenant à diverses catégories lexicales (nom ou syntagme nominal, adjectif ou syntagme adjectival, verbe ou locution verbale et adverbe ou locution adverbiale). Nous nous proposons donc, dans cette communication, d'illustrer les diverses considérations théoriques et constats présentés ci-dessus par l'étude de cas issus de nos travaux, en les mettant au besoin en miroir avec des exemples provenant d'autres domaines de spécialité.

#### **Références bibliographiques**

- Ait Taleb, S. (1993). Rapport de la phraséologie avec la terminologie. *Terminologies nouvelles*, 10, 13-15.
- Depecker, L. (1999). Monème, syntème et phrasème : Essai d'introduction du concept de phrasème dans la théorie fonctionnaliste. *La Linguistique*, 35(2), 23-46.
- Gautier, L. (2000). Du terme au phrasème en droit constitutionnel : un double aller-retour ? In Gréciano, G. et al. (éds.), *Micro- et macroléxèmes et leur figement discursif : actes du colloque international ; 6-7-8 décembre 1998, Saverne*. Louvain : Peeters, 285-296.
- Gautier, L. (2002). Terme, phraséoterm, phrasème : questions de délimitation en langue spécialisée. In *Le Continuum en linguistique*. Tunisie : Faculté des Lettres et Sciences Humaines de Sousse. 153-172.
- Gréciano, G. (1993). Vers une modélisation phraséologique : Acquis et projets d'EUROPHRAS. *Terminologies nouvelles*, 10, 16-22.
- L'Homme, M.-C. (1998). Le statut du verbe en langue de spécialité et sa description lexicographique. *Cahiers de lexicologie : Revue internationale de lexicologie et lexicographie*, 73(2), 61-84.
- Mautner, G. et al. (éds.), 2017. *Handbook of business communication : linguistic approaches*. Boston Berlin : De Gruyter Mouton.
- Resche, C. (1997). Prolégomènes à la phraséologie comparée en langue de spécialité : exemple de l'anglais et du français de la finance. *ASp. la revue du GERAS*, 15-18, 487-503.
- Rousseau, D.-A. (2019). *Un cas de résurgence terminologique : La terminologie musicale en usage en France et en Angleterre à la seconde moitié du XVII<sup>e</sup> siècle*. Thèse doctorale en Lexicologie et Terminologie multilingues, Traduction (LTMT). Université Lumière Lyon 2.
- Silva, R. et al. (2004). Entre langue générale et langue de spécialité une question de collocations. *Ela. Études de linguistique appliquée*, 135(3), 347-359.

### **« Ajustement cosmétique », « don manuel », « rescrit social » et bien d'autres : le phraséoterm à la confluence de la langue naturelle, de la langue de spécialité et des néoformations**

Michela Tonti (Università di Bologna DIT - Forlì)  
Pietro Moretti

Les besoins de nomination des parlants relevant d'un domaine spécialisé se traduisent par l'acte de nommer à partir d'une terminologie qui n'est qu'une dimension d'une plus riche langue de spécialité, une partie insécable de la langue naturelle, inconcevable sans celle-ci. Cependant, le stock des phrasèmes disponibles nécessite de nouveaux termes pour pouvoir répondre aux exigences croissantes et mouvantes de locuteurs, certes avisés et expérimentés, mais qui ont tout de même besoin de mettre un terme sur un concept nouveau. Renouant avec Gréciano (1997) et Rosenbaum Franková (2016), on s'intéressera à la terminologie polylexicale, en examinant l'articulation entre langue de spécialité et langue générale d'un côté, nouvelles créations terminologiques de l'autre ; on observera notamment si un modèle de néologie terminologique de prédilection s'impose (Humbley, 2018). Le cadre retenu est celui de la phraséologie scientifique afférant au domaine de la comptabilité puisant dans un corpus constitué par nos soins.

Dans un premier temps, il nous appartiendra de définir le corpus expérimental que nous avons bâti à partir de la base de données *ProQuest* portant sur un macro-domaine qui a trait à l'économie d'entreprise. Le matériel réuni s'échelonne sur un laps de temps qui va de 1999 à 2016, ce qui constitue un atout aussi bien du point de vue de la nouveauté de ses observables que de son implantation car on ne peut pas « jug[er] de l'implantation d'un terme en dehors d'une certaine diachronie » (Depecker, 1997 : 9-10). Ce corpus ayant une valeur d'échantillonnage des néoformations du domaine que nous explorons est de dimensions suffisamment considérables pour sa représentativité. Il comporte 700000 mots et se compose d'un florilège de textes techniques : revues scientifiques, rapports d'entreprise, guides opérationnels, ouvrages

spécialisés, feuilles de travail en audit interne. Le corpus sera exploité à partir de la version 3.0 de l'outil d'acquisition automatique *TermoStat* (Drouin, 2003 et 2004). Une liste de candidats-termes sera dressée au moyen de ce logiciel, appliquant la technique mise au point par Drouin qui porte sur la comparaison d'un corpus de référence de langue générale réunissant 560000 formes différentes issues du quotidien français *Le Monde* (2002) et notre corpus de langue de spécialité. Au vu de notre perspective comparatiste, nous nous servons de notre liste de candidats-termes en guise de base empirique pour détecter et identifier l'attestation de ces candidats-termes dans un deuxième corpus de référence de grande envergure *FrTenTen17* (conçu en 2017, 10 milliards de tokens environ).

Les résultats de l'extraction feront premièrement l'objet d'une opération de tamisage visant à cerner une réelle saillance terminologique et conceptuelle des données réunies : à cet effet, la distribution quantitative fournit un critère de sélection (Sinclair, 1991 : 70). Deuxièmement, ils seront classés en fonction de leur attestation dans une sélection de bases de données terminologiques existantes : *IATE*, *BelTerme* (Belgique), *Lingua-PC online* (Suisse – Chancellerie d'État du canton de Berne), *TERMIUM Plus* (Canada). Cette fouille outillée nous permet de cerner des néoformations intéressantes du point de vue terminologique qui seront catégorisées et analysées en fonction des modèles de néologie terminologique recensés par Humbley (2018) et des critères syntactico-sémantiques et lexicaux permettant de distinguer plus clairement les expressions de la langue générale des candidats-néologismes relevant de la langue de spécialité. Des approfondissements statistiques seront également effectués.

Nous consacrons notre troisième partie à l'analyse d'études de cas qui s'avèrent des phraséotermes nouveaux n'apparaissant dans aucune des bases de données susmentionnées. Nous allons procéder à une classification qui prendra en compte les termes monolexicaux, à savoir des noms génériques qui restent des outils d'approximation comme « ajustement », « don », que seule la spécification par collocataires (Gréciano, 1997 : 34) rend contextuellement efficaces. En guise d'exemple, nous mentionnons « ajustement cosmétique », qui a droit de cité aussi bien dans le domaine du droit que de la finance, alors que le nouveau collocataire responsable de la phraséologisation du terme « don » est « manuel », qui va sémantiser une donation bien précise. Nous allons démontrer que si la langue générale semble l'emporter là où un nouveau besoin de nomination s'impose, il n'en demeure pas moins vrai que parmi les néoformations des phraséotermes spécialisés détectés, certains d'entre eux sont le produit de formatifs exclusivement savants comme « rescrit social », n'apparaissant dans aucune ressource terminologique avec ce collocataire. Ce dernier joue pleinement son rôle spécifique alors que seul « fiscal » est retenu dans nos supports terminologiques (*IATE* et *TERMIUM Plus*). L'originalité conceptuelle des néoformations détectées constituera un critère supplémentaire de sélection pour l'analyse des études de cas dont nous observerons les propriétés syntaxiques et sémantiques à la confluence avec la langue naturelle.

## Références bibliographiques

btb.termiumplus.gc.ca (*TERMIUM*).

Depecker, L. (éd.) (1997). *La mesure des mots : cinq études d'implantation terminologique*. Rouen : Publications de l'Université de Rouen.

Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, 9(1), 99-117.

Drouin, P. (2004). Detection of domain specific terminology using corpora comparison. In Lino, M.T. et al. (éds.) *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 04)*. Lisbonne : European Language Resources Association (ELRA).

franca.cfwb.be/bd/bd.htm (*BelTerme*)

*French Web 2017 (frTenTen17)*, <https://www.sketchengine.eu/>

Gréciano, G. (1997). Collocations rythmologiques. *Meta*, 42(1), 33-44.

<http://termostat.ling.umontreal.ca/> (*TermoStat*)

Humbley, J. (2018). *La néologie terminologique*. Limoges : Lambert-Lucas.

iate.europa.eu (*IATE*).

linguapc.apps.be.ch (*Lingua-PC*).

Rosenbaum Franková, L. (2016). Phrasèmes spécialisés dans les textes économiques. *Cahiers de lexicologie*, 108(1), 43-57.

search-proquest-com (*ProQuest*)

Sinclair, J. (1991). *Corpus, concordance, collocations*. Oxford : Oxford University Press.

## Quels traitements sémantiques pour la phraséologie scientifique transdisciplinaire dans une perspective d'aide à la rédaction ?

Agnès Tutin (Université Grenoble Alpes)

La phraséologie scientifique transdisciplinaire est l'ensemble des expressions polylexicales qui ne renvoient pas aux objets d'étude des disciplines scientifiques, mais plutôt aux procédures, méthodes et éléments du raisonnement propres au genre des écrits scientifiques. Ce lexique transversal aux différentes disciplines (même si certaines expressions sont particulièrement fréquentes dans certaines familles de discipline) porte ainsi sur plusieurs types d'éléments comme les suivants :

- L'activité scientifique et l'évaluation : *collecter des données, mener une expérimentation, résultats encourageants.*
- Le raisonnement scientifique : *nous arrivons à la conclusion, c'est pourquoi ...*

- Le métadiscours et le métatexte : *dans un premier temps, comme on l'a vu, au contraire.*

Bien entendu, cette phraséologie présente des points de recoupement avec la terminologie (l'expression *analyse du discours* relèvera probablement de la terminologie en linguistique, alors qu'elle s'apparentera à un élément de méthode en sociologie ou en sciences politiques).

Sur le plan structurel, cette phraséologie est assez diverse, puisqu'elle comporte plusieurs types d'expressions, en particulier :

- Des locutions complètement figées, incluant des mots grammaticaux, comme *point de vue, mettre en œuvre, c'est-à-dire ...*
- Des collocations, qu'on définira comme des associations binaires privilégiées : *jouer un rôle, tester empiriquement, étroite collaboration ...*
- Des routines sémantico-rhétoriques comme *il est important de souligner, comme on l'a vu/noté ...*

Des ressources phraséologiques de locutions et collocations ont été constituées dans notre équipe pour le français<sup>1</sup> (Jacques & Tutin, 2018 ; Hatier *et al.*, 2016), et des études sont en cours pour le repérage des routines sémantico-rhétoriques (Ji & Tutin, 2019). Une description syntaxique et sémantique simple de ces éléments a été proposée, mais un accès à la fois plus structuré et plus intuitif à ces données reste à effectuer pour proposer un outil d'aide à la rédaction pour les apprentis chercheurs, où le rôle des corpus serait central. Des outils de ce type ont été proposés pour l'anglais (Granger & Paquot, 2010) ainsi que pour l'espagnol (Alonso-Ramos *et al.*, 2017), mais le français dispose surtout à l'heure actuelle pour le discours scientifique de ressources où la phraséologie transdisciplinaire est abordée en complément de la terminologie (Pecman & Kübler, 2011, Pecman 2018)<sup>2</sup>.

Dans cette communication, nous souhaitons proposer une réflexion sur l'organisation d'une base de données phraséologiques à destination des apprentis scripteurs organisée autour des principes suivants : une organisation onomasiologique, basée sur un métalangage simple, exploitant des mécanismes analogiques et accordant une place centrale aux exemples sur corpus.

## Références bibliographiques

- Alonso-Ramos, M., García-Salido, M., Garcia, M. (2017). Exploiting a corpus to compile a lexical resource for academic writing: Spanish lexical combinations. In *Electronic Lexicography in the 21st Century. Proceedings of eLex 2017 Conference*, 571-586.
- Granger, S., Paquot, M. (2010). The Louvain EAP Dictionary. In *Proceedings of the XIV EURALEX International Congress. Leeuwarden/Ljouwert: Fryske Akademy*, 321-326.
- Hatier, S., Augustyn, M., Yan, R., Tran, T. T. H., Tutin, A., & Jacques, M. - P. French cross-disciplinary scientific lexicon: extraction and linguistic analysis. In T. Margalitadze & G. Meladze (éds.), *Proceedings of the XVII EURALEX International congress Lexicography & Linguistic diversity*, 355-365.
- Jacques, M.-P., Tutin, A. (2018) (éds). *D'une discipline à l'autre : lexicque transversal et formules discursives des sciences humaines*. ISTE.
- Ji, Y., Tutin, A. (2019). Les routines métalinguistiques dans les écrits scientifiques en français. *Studii de lingvistică*, Les expressions métadiscursives dans les langues romanes : aspects syntaxiques, pragmatiques et sociolinguistiques, 9 (2), 177-200.
- Pecman, M., Kübler, N. (2011). ARTES : an online lexical database for research and teaching in specialized translation and communication. In *Proceedings of the First International Workshop on Lexical Resources*.
- Pecman, M. (2018). *Langue et construction de connaisSENSes. Energie lexico-discursive et potentiel sémiotique des sciences*. L'Harmattan.

## Vers une méthodologie pour l'extraction automatique des collocations en terminologie médicale

Federica Vezzani (Università degli Studi di Padova)

### 1. Introduction

Cette étude s'inscrit dans le contexte du traitement automatique des langues (TAL) et vise à proposer une méthodologie efficace pour l'extraction automatique des collocations à collocatif verbal relevant du langage médical.

Il existe un grand nombre d'études sur les unités phraséologiques, de type locution et/ou collocation, en langue de spécialité : (Galinski 1990 ; Rousseau 1993 ; Candel 1995 ; Rosenbaum 2016) pour le langage économique ; (Tutin et Grossmann 2014 ; Jacques et Tutin 2018), en particulier, pour le langage scientifique. Dans ce contexte, l'axe de recherche de la « Phraséologie Computationnelle » prend forme et rassemble les intérêts de nombreux chercheurs de TAL en ce qui concerne l'écémage, l'identification et l'extraction (semi-) automatique d'unités phraséologiques. Pour ne citer que les contributions les plus récentes, voir le chapitre de (Heid 2008), l'enquête menée par (Constant *et al.* 2017) sur le traitement

<sup>1</sup> Voir la ressource (version bêta) sur: [lst.demarre-shs.fr](http://lst.demarre-shs.fr)

<sup>2</sup> La base ARTES intègre plusieurs langues, dont le français.

des expressions polylexicales, le livre rassemblant les actes du colloque Europhras 2019<sup>1</sup> de (Pastor et Birukou 2019) et, enfin, le livre récent intitulé *Computational Phraseology* et édité par (Pastor et Colson 2020). Le processus d'extraction automatique est principalement basé sur des approches linguistiques, statistiques et/ou hybrides dont les mesures ont été importées par le domaine de la recherche d'information (RI) : *Term frequency/Inverse Document Frequency* (TF/IDF) (Salton et Yang 1973), *Information Mutuelle* (Church et Hanks 1990), *T-Score* (Church et al. 1991), *C/NC value* (Frantzi et al. 1998). Enfin, des ressources spécifiquement conçues pour cette tâche ont été réalisées afin d'augmenter les performances d'extraction, par exemple : *TermoStat* (Drouin 2003), *BiTermEx* (Planas 2012) et *TermEvaluator* (Inkpen et al. 2016), etc.

## 2. Méthodologie d'extraction automatique

L'objectif de notre étude est de proposer une méthodologie pour l'identification et l'extraction automatique des collocations à collocatif verbal dans des textes médicaux. Reprenant la définition fournie par (Polguère 2015), nous considérons la collocation comme une structure binaire, formée d'une base et d'un collocatif, constituant une association lexicale privilégiée dont le sens est compositionnel. Notre méthodologie consistera donc, tout d'abord, à identifier la base de la collocation constituée de noms (dans le cas spécifique, les termes véhiculant une signification médicale) et, ensuite, les collocatifs verbaux associés.

Le corpus choisi comprendra tous les textes rédigés en français contenus dans le Manuel MSD disponible en ligne.<sup>2</sup> En perspective diastatique et multilingue, cette ressource fournit des informations spécialisées à l'usage des professionnels et des articles de vulgarisation pour consultation du grand public. Une fois sélectionné notre corpus, nous procéderons à l'étiquetage morphosyntaxique des textes en utilisant un logiciel implémenté en langage de programmation R<sup>3</sup> afin d'attribuer une étiquette morphosyntaxique à chaque mot contenu dans notre corpus. Après avoir rassemblé le corpus et étiqueté les textes, nous procéderons avec l'extraction automatique des termes médicaux de type nom. Pour cette dernière tâche, nous utiliserons tous les termes médicaux en français contenus dans le thésaurus MeSH *terms*.<sup>4</sup> Ce dictionnaire médical, généralement utilisé pour l'indexation d'articles biomédicaux, jouera le rôle de ressource de correspondance pour l'extraction de tous les termes médicaux dans notre corpus. Ensuite, nous mènerons une analyse de dépendance en utilisant les corpus arborés (*treebank*) fournis dans le cadre du projet *Universal Dependencies* (UD),<sup>5</sup> toujours à l'aide du langage de programmation R (avec le package *UDpipe*),<sup>6</sup> afin d'identifier automatiquement tous les verbes associés aux noms précédemment extraits. Cette méthodologie nous permettra, donc, d'extraire automatiquement les collocatifs verbaux associés aux bases des collocations sous la forme de noms : en considérant le terme « médicament », par exemple, nous nous attendons à extraire automatiquement tous les verbes associés au substantif afin d'en étudier les occurrences et en vérifier les associations préférentielles comme « prescrire un médicament », « administrer un médicament », « tolérer un médicament », etc.

## 3. Conclusions et perspectives

Dans cette étude, nous proposons une méthodologie pour l'extraction automatique des unités phraséologiques dans le langage médical. En particulier, nous nous concentrerons sur l'identification des collocations à collocatif verbal. À partir de cette analyse et compte tenu de la variété diastatique du lexique médical, c'est-à-dire de la terminologie utilisée entre les professionnels de la santé et de la terminologie destinée aux profanes (Vecchiato et Gerolimich 2013), nous visons à mener une étude contrastive future afin de vérifier s'il existe des différences dans la représentation du phénomène de la collocation dans les deux variations de registre.

## Références bibliographiques

- Candel, D. (1995). Locutions en langues de spécialité. *Cahiers du français contemporain*, 2, 151-173.
- Church, K.W. and Hanks P. (1990). Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, 16 (1): 22-29.
- Church, K., Gale W., Hanks P., Hindle D. (1991). Using Statistics in Lexical Analysis. *Lexical Acquisition : Exploiting on-Line Resources to Build a Lexicon*, 115-164.
- Constant, M., Eryigit, G., Monti, J., Van Der Plas, L., Ramisch, C., Rosner, M., & Todirascu, A. (2017). Multiword expression processing : A survey. *Computational Linguistics*, 43(4), 837-892.
- Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, 9(1), 99-115.
- Frantzi, K. T., Ananiadou S., and Tsujii J. (1998). The C-Value/Nc-Value Method of Automatic Recognition for Multi-Word Terms. In *International Conference on Theory and Practice of Digital Libraries*, 585-604. Springer.
- Galinski, C. (1990). Terminology and phraseology. Terminology Science and Research. In *Journal of the International Institute for Terminological Research* (IITF), 1.
- Jacques, M. P., & Tutin, A. (2018). *Lexique transversal et formules discursives des sciences humaines*. ISTE Group.
- Heid, U. (2008). Computational phraseology. An overview. In Granger, S., Meunier, F. (eds.), *Phraseology: An Interdisciplinary Perspective*, 337-360.

<sup>1</sup> <http://www.lexytrad.es/europhras2019/>.

<sup>2</sup> <https://www.msmanuals.com/fr/accueil>.

<sup>3</sup> <https://cran.r-project.org/web/packages/udpipe/index.html>.

<sup>4</sup> <https://meshb.nlm.nih.gov/search>.

<sup>5</sup> <https://universaldependencies.org>.

<sup>6</sup> <https://cran.r-project.org/web/packages/udpipe/index.html>.

- Inkpen, D., Paribakht, T. S., Faez, F., & Amjadian, E. (2016). Term Evaluator : A Tool for Terminology Annotation and Evaluation. *Int. J. Comput. Linguistics Appl.*, 7(2), 145-165.
- Pastor, C., & Birukou. (2019). *Computational and Corpus-Based Phraseology*. Springer International Publishing.
- Pastor, G. C., Colson, J. P. (2020). *Computational Phraseology*. John Benjamins Publishing Company.
- Planas, E. (2012). BiTermEx Un prototype d'extraction de mots composés à partir de documents comparables via la méthode compositionnelle. In *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012*, 2, TALN.
- Polguère, A. (2015). Non compositionnalité : ce sont toujours les locutions faibles qui trinquent, *Verbum*, 37(2), 257-280.
- Rosenbaum Franková, L. (2016). Phrasèmes spécialisés dans les textes économiques. *Cahiers de lexicologie*, 108, 43-57.
- Rousseau, L. J. (1993). Terminologie et phraséologie, deux composantes indissociables des langues de spécialités. *Terminologies nouvelles—Phraséologie Actes du séminaire international*, 10, 9-11.
- Salton, G., and Chung-Shu Y. (1973). On the Specification of Term Values in Automatic Indexing. *Journal of Documentation*, 29(4), 351–372.
- Tutin, A., Grossmann F. (éds.). (2014). *L'écrit scientifique : du lexique au discours*. Presses universitaires de Rennes.
- Vecchiato, S., Gerolimich, S. (2013). La langue médicale est-elle « trop complexe » ? *Nouvelles perspectives en sciences sociales : revue internationale de systémique complexe et d'études relationnelles*, 9(1), 81-122.