



**HAL**  
open science

# PHRASEOROM. Phraséologie du roman contemporain français, anglais et allemand

Iva Novakova

► **To cite this version:**

Iva Novakova. PHRASEOROM. Phraséologie du roman contemporain français, anglais et allemand.  
[Rapport de recherche] ANR. 2020. hal-02524894

**HAL Id: hal-02524894**

<https://hal.univ-grenoble-alpes.fr/hal-02524894v1>

Submitted on 1 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Projet ANR-15-FRAL-0009**

**PHRASEOROM**

Programme FRAL 2015

## A IDENTIFICATION

Acronyme du projet	PHRASEOROM
Titre du projet	Phraséologie du roman contemporain français, anglais et allemand
Coordinateur du projet (société/organisme)	Laboratoire Lidilem E.A. 609, Université Grenoble Alpes UGA)
Date de début du projet	01.05.2016
Date de fin du projet	31.12.2019
Labels et correspondants des pôles de compétitivité (pôle, nom et courriel du corresp.)	
Site web du projet, le cas échéant	<a href="https://phraseorom.univ-grenoble-alpes.fr/">https://phraseorom.univ-grenoble-alpes.fr/</a>

Rédacteur de ce rapport	
Civilité, prénom, nom	Mme Iva NOVAKOVA
Téléphone	
Adresse électronique	
Date de rédaction	31.01.2020

Liste des partenaires présents à la fin du projet (société/organisme et responsable scientifique)	Les équipes de l'UGA (Lidilem et Litt&Arts), D. Legallois (Paris 3 Sorbonne nouvelle), S. Diwersy et C. Jacquot (U. de Montpellier), les partenaires allemands : U. d'Osnabrück (Prof. D. Siepmann), de l'U. de Bonn (Prof. M. Gymnich), de l'U. d'Erlangen (Prof. L. Fesenmeier)
---	---

## B RESUME CONSOLIDE PUBLIC

### B.1 RESUME CONSOLIDE PUBLIC EN FRANÇAIS

*La phraséologie du roman contemporain français, anglais et allemand*

1. **Vers une analyse des genres littéraires contemporains, à travers leurs expressions lexico-syntaxiques spécifiques**

Le principal objectif du projet PhraseoRom <https://phraseorom.univ-grenoble-alpes.fr/> a été d'élaborer une typologie structurelle et fonctionnelle des constructions lexico-syntaxiques (CLS) spécifiques au discours romanesque francophone, anglophone et germanophone à partir des années 1950. Nous avons exploré, à partir de grands corpus littéraires trilingues et avec des méthodes lexicométriques, la pertinence des phénomènes phraséologiques, et en particulier du *motif textuel*, pour la caractérisation de la langue littéraire. Les œuvres constitutives de ces corpus ont été réparties en 6 sous-genres : littérature générale (ou « blanche »), romans sentimentaux, policiers, historiques, de science-fiction, fantasy. L'enjeu a été d'identifier et de décrire les genres littéraires non pas en termes de rhétorique ou de style individuel des auteurs, comme c'est souvent le cas en stylistique classique, mais à travers des *patrons lexico-syntaxiques récurrents* extraits automatiquement avec des outils du Traitement automatique du langage (TAL). Plus généralement, il a été montré que la linguistique de corpus aide à découvrir les stéréotypes langagiers généraux qui sous-tendent le genre romanesque.

## 2 Vers une approche interdisciplinaire pour différencier les sous-genres littéraires contemporains

Les méthodes lexicométriques ont permis d'extraire et de regrouper automatiquement plusieurs milliers de constructions lexico-syntaxique, à partir des corpus littéraires trilingues. Les récurrences significatives de ces expressions ont été établies par rapport à un corpus de contraste (textes journalistique, scientifiques). Un système original d'annotation sémantique et stylistique de ces expressions a été élaboré. L'analyse des données sur les plans sémantique et syntaxique a été systématiquement articulée à une analyse discursive stylistique dans un but comparatif (inter-genres et inter-langues). Le motif n'est plus un élément fictionnel symbolique ou constitutif de l'imaginaire d'une œuvre, mais une *observable phraséologique*, composée d'unités continues ou discontinues combinant plusieurs éléments: formes de mots, catégories morpho-syntaxiques, mots-outils. Les résultats de ce projet interdisciplinaire en Humanités numériques auront des retombées en linguistique et en stylistique contrastives, en traductologie, en écriture créative.

### Résultats majeurs du projet.

Le projet a révélé la pertinence de la notion de *motif textuel* pour la caractérisation des genres littéraires et a permis de l' « exporter » vers la communauté de chercheurs anglo-saxons en Humanités numériques. Il a complété l'étude des genres littéraires par une approche originale fondée sur la linguistique de corpus. Il a contribué à la réflexion méthodologique sur l'émergence de la stylistique « outillée » (numérique). Des contacts scientifiques ont été établis avec des chercheurs travaillant sur des problématiques proches.

### Production scientifique

Le projet a réuni 32 participants de 10 universités. Il a donné lieu à plus de 90 communications, publications, conférences invitées. Parmi les faits marquants, on citera le Colloque *Phraseology and Stylistics of Literary language*, en mars 2019 à l'Université d'Erlangen-Nuremberg (45 participants de 15 pays) et l'ouvrage *Phraseology and Style in Subgenres of the Novel*, Palgrave MacMillan. La *PhraseoBase*, réunissant les corpus, les manuels méthodologiques et les données annotées du projet a été inaugurée en décembre 2019.



Le projet PhraseoRom est un projet de recherche fondamentale coordonné par Iva Novakova (Professeur à l'Université Grenoble Alpes, UGA, France) et Dirk Siepmann Professeur à l'Université d'Osnabrück, Allemagne). Il associe des chercheurs de l'UGA (Laboratoires Lidilem, et Litt&Arts), de l'U. Paul Valéry Montpellier 3 et de Paris 3 Sorbonne Nouvelle, des chercheurs de 3 universités allemandes (Osnabrück, Bonn, Erlangen) ainsi que des collaborateurs extérieurs de Pologne et de Suisse. Le projet a commencé en janvier 2016 et a duré 48 mois. Il a bénéficié d'une aide ANR de 325 208 € pour un coût global de l'ordre de 717 758 €.

## **B.2 RESUME CONSOLIDE PUBLIC EN ANGLAIS**

### *Phraseology in the English, French and German novel*

#### *1. Towards an analysis of contemporary literary genres, through their specific lexico-syntactic expressions*

The chief goal of the PhraseoRom project has been to elaborate a structural and functional typology of lexico-syntactic constructions specific to English, French and German novelistic discourse. Drawing on large trilingual corpora of literary language, we have used corpus-linguistic (or, more precisely, lexicometric) methods to explore the usefulness and relevance of the concept of phraseology, and more particularly, the notion of 'motif' to the description of literary language. The writings that make up the corpus were divided into six sub-genres : general, sentimental, crime, historical, fantasy and science fiction. The aim was to give an account of genre in terms of corpus-extracted recurrent lexico-syntactic patterns rather than in the traditional manner of stylistics, which has tended to focus on rhetoric and individual style. This approach has been shown to be of great benefit in uncovering the linguistic and cognitive stereotypes underlying the genre of the novel.

#### *2. Towards an interdisciplinary approach for differentiating contemporary fiction subgenres*

Our lexicometric methodology has allowed us to a) automatically extract from our trilingual literary corpora several thousand constructions and b) to sort these according to predefined criteria. A comparable corpus of newspaper and academic language was used to determine the significance of recurrent constructions. We have also developed a systematic new

approach to annotating the expressions in question both semantically and stylistically. The syntactic-semantic analysis of the data was systematically combined with a discursive and stylistic analysis aiming at cross-genre and cross-linguistic comparison. The motif thus becomes not so much a fictional element imbued with literary symbolism or contributing to the construction of a fictional universe, but rather an *observable phraseological phenomenon* made up of either continuous or discontinuous units that may combine various types of elements : word forms, morpho-syntactic categories, function words. The results of this interdisciplinary project have implications for linguistics and contrastive stylistics, translation studies and creative writing.

### ***Main results of the project***

The project has shown that the notion of ‘textual motif’ can be fruitfully applied to the description of literary genres and has allowed us to introduce this new concept into Anglo-American discourse on the Digital Humanities. It extends the traditional armentarium of literary studies through a new corpus-linguistic methodology. It has contributed to current debate and reflection on methodological and theoretical issues surrounding corpus stylistics. Contacts have been established with researchers engaged with related concerns.

### ***Publications***

The project, which brought together 32 participants from ten universities, has given rise to more than 90 papers, publications, and invited lectures. Highlights include a symposium on the Phraseology and Stylistics of Literary Language, held at the University of Erlangen-Nuremberg in March 2019 (45 participants from 15 countries) and the book *Phraseology and Style in Subgenres of the Novel*, Palgrave-Macmillan. The PhraseoBase, a web-based collection of our annotated corpora, methodological manuals and data, was inaugurated in December 2019.



The PhraseoRom project, led by Iva Novakova (Professor at the University of Grenoble Alpes, UGA, France) and Dirk Siepmann (Professor at the University of Osnabrück, Germany), has focussed on basic rather than applied research. It has brought together researchers from UGA (Laboratoires Lidilem, and Litt&Arts), U. Paul Valéry Montpellier 3 and Paris 3 Sorbonne Nouvelle), as well as researchers from three German universities (Osnabrück, Bonn, Erlangen) and external partners from Poland and Switzerland. The project started in January 2016 and ran for 48 months. It received an ANR grant of €325 208, with total project costs amounting to around €717 758.

## C MEMOIRE SCIENTIFIQUE (NON CONFIDENTIEL)

### C.1 RESUME DU MEMOIRE

Le principal objectif du projet PhraseoRom <https://phraseorom.univ-grenoble-alpes.fr/> a été d'élaborer une typologie structurelle et fonctionnelle des constructions lexico-syntaxiques spécifiques au discours romanesque francophone, anglophone et germanophone à partir des années 1950. Nous avons exploré, en nous appuyant sur de grands corpus littéraires trilingues et avec des méthodes lexicométriques, la pertinence des phénomènes phraséologiques et, en particulier, du *motif textuel* pour la caractérisation de la langue et des sous-genres littéraires (littérature générale, romans sentimentaux, policiers, historiques, de science-fiction, fantasy). L'enjeu a été d'identifier et de décrire les genres littéraires non pas en termes de rhétorique ou de style individuel des auteurs, comme c'est souvent le cas en stylistique classique, mais à travers des patrons lexico-syntaxiques récurrents extraits automatiquement avec des outils du Traitement automatique du langage (TAL).

### C.2 ENJEUX ET PROBLEMATIQUE, ETAT DE L'ART

Notre principale question de recherche a été de savoir *si* et *comment* la phraséologie étendue permet-elle de définir la langue et les genres littéraires contemporains? Le projet a été construit autour de trois principaux volets: 1. *Corpus et TAL* ; 2. *Linguistique* ; 3. *Stylistique*.

#### 2.1. Le volet Corpus et Traitement automatique du langage (TAL) (responsables O.

Kraif, UGA et S. Diwersy, U. Paul Valéry, Montpellier 3)

Il a été prévu dans ce volet d'enrichir les corpus littéraires français, anglais et allemands, collectés lors d'un précédent projet AND DFG Emolex (2009-2013), ainsi que de développer des scripts plus performants pour l'extraction d'unités polylexicales à partir de ces corpus. Les applications récentes en TAL permettent désormais d'aller au-delà du traitement de simples *n-grammes* de mots-formes. Une approche prometteuse a été adoptée par Quiniou *et al.* (2012), qui exploitent les séquences en surface au moyen d'une technique connue en fouille de données : l'extraction de *motifs séquentiels*. Une autre approche innovante, qui va au-delà de la simple séquence de mots en surface, a été proposée par les membres de PhraseoRom : Kraif & Diwersy (2012) et Kraif (2016). Appliquant une technique d'extraction itérative, ils ont développé une méthode d'extraction, sans précalcul, des expressions récurrentes autour d'un mot pivot donné. Ces expressions représentent des sous-arbres syntaxiques (Figure 1) :

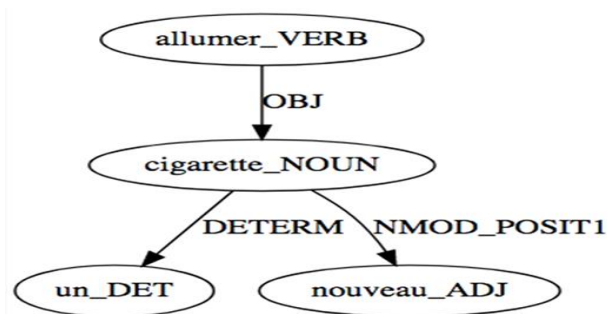


Figure 1 : Extraction automatique d'un *arbre lexico-syntaxique récurrent (ALR)* correspondant à l'expression *allumer une nouvelle cigarette*, spécifique au sous-genre policier (POL)

A la différence de Quiniou *et al.* (2012), les objets extraits dans le cadre du projet PhraseoRom correspondent à des *structures hiérarchiques* (arbres de dépendance).

**2.2. Le volet linguistique:** Responsables : Iva Novakova (UGA) et Dirk Siepmann (U. d'Osnabrück)

Le point de départ du projet a été le constat selon lequel la langue romanesque se caractérise par la surreprésentation statistiquement significative de certains phénomènes linguistiques (lexèmes, mots clés, collocations, colligations, schémas actanciels, phraséologismes) (Siepmann, 2015). Une amorce de validation de cette hypothèse existait déjà dans l'étude pionnière sur l'anglais de Stubbs & Barth (2003) mais elle n'a été validée que sur des corpus de taille très modestes, excluant les unités « lexicalement riches », immédiatement identifiables comme spécifiques au genre littéraire comme par ex. : *His thoughts were interrupted by* (ex. de Siepmann 2015). Par ailleurs, Biber et al. (1999) ont aussi identifié certaines propriétés grammaticales spécifiquement littéraires (les constructions *that bloody car of mine* ou en *there*) mais leurs analyses ne prennent pas en compte les mots-clés, ni les segments répétés.

L'enjeu, lié à ce contexte scientifique, a été d'annoter sur le plan syntaxico-sémantique des constructions spécifiques, extraites de nos corpus trilingues sous formes d'ALR (cf. figure 1). Dans les objectifs initiaux du projet, il s'agissait de comparer la langue littéraire des romans contemporains par rapport à un corpus de contraste (textes scientifiques, journalistiques, TED talks, sous-titres de films). Très rapidement, dès la 1<sup>ère</sup> année du projet, les membres des équipes ont saisi l'intérêt d'élargir la réflexion phraséologique à la distinction non pas seulement de la langue littéraire en général mais aussi des sous-genres constitutifs de nos corpus.

Par ailleurs, au fur et à mesure de l'avancement du projet, il s'est avéré pertinent de regrouper les constructions lexico-syntaxiques spécifiques, extraites de nos corpus, en *motifs textuels*. Cette notion récente en phraséologie étendue, introduite par D. Legallois (2006), membre de PhraseoRom et d'autres collègues (Longrée & Mellet 2013), correspond à des patrons lexico-syntaxiques récurrents qui présentent des régularités et des variations lexicales et syntaxiques (au niveau paradigmatique et syntagmatique), établies avec des méthodes lexicométriques, ainsi que des fonctions discursives (Adam 2011) spécifiques. De par son caractère à la fois structurant et caractérisant (Longrée & Mellet 2013), le motif s'est révélé comme étant une notion opératoire pour nos objectifs, liés à la distinction des sous-genres littéraires.

**2.3. Le volet stylistique:** responsables Laetitia Gonon (UGA) et M. Gymnich (U. de Bonn)

Il existe un grand nombre de travaux sur le style de différents auteurs classiques (p. ex. Philippe & Piat, 2009 sur Flaubert, Zola, Proust). Il a été également montré (*idem*) qu'une langue « autre » avait émergé en contexte littéraire en France autour de 1850. Par ailleurs, les critères retenus dans les théories des genres ont été presque exclusivement basés sur leur contenu (Gymnich, Neumann & Nünning 2007). Il n'existait donc pas de travaux de grande ampleur sur le roman contemporain dans une perspective phraséologique et contrastive (inter-genres et inter-langues). Ainsi, l'approche des genres par le contenu demandait à être complétée par une approche fondée sur les outils de la linguistique de corpus, permettant d'extraire de grandes masses de données pour l'analyse stylistique. Le principal objectif du volet stylistique a été d'identifier et d'analyser les *fonctions discursives* (FD) des motifs textuels repérés par les linguistes.

### C.3 APPROCHE SCIENTIFIQUE ET TECHNIQUE

L'analyse des constructions polylexicales spécifiques qui génèrent des motifs textuels s'est avérée une problématique beaucoup plus complexe, comparée à l'étude des collocations binaires (par ex. les collocations d'émotions étudiées dans le cadre du projet ANR DFG Emolex). Cet objet d'étude plus complexe a donc nécessité le recours à plusieurs approches scientifiques. Pour l'extraction des données ont été appliquées des approches *lexicométriques* (Lebart et al. 1998), ainsi que les *Keywords analysis* (Scott & Tribble 2006). Nos observations ont été essentiellement guidées par les données (l'approche inductive *corpus driven*). Les unités linguistiques ont été, par la suite, analysées sur 4 niveaux : lexical, sémantique, syntaxique et discursif dans la continuité des approches globales, fondées sur les *modèles fonctionnels et contextualistes* (Sinclair 2004, Hoey 2005). Enfin, nous avons comparé certains traits spécifiques de la langue littéraire anglaise et française (et ponctuellement allemande). Le véritable apport du projet a été la mise en place de la méthodologie pour l'extraction et le regroupement des ALR en motifs textuels. La spécificité statistique des ALR pour tel ou tel sous-genre a été établie selon le calcul du rapport de vraisemblance *log likelihood ratio* (LLR).

### C.4 RESULTATS OBTENUS

#### 4.1. Volet corpus et TAL

Les objectifs concernant l'enrichissement des corpus littéraires comparables ont été largement atteints, voire dépassés. Pour chacune des trois langues du projet, plus de 1000 textes ont été collectés pour un total de 110 M. de mots. Des corpus parallèles (en-fr et fr-en) ont été également constitués pour un total de 100 M. de mots. Les textes ont été répartis en 6 sous-genres avec l'aide de nos collègues stylisticiens: littérature générale (GEN), romans policiers (POL), sentimentaux (SENT), historiques (HIST), science-fiction (SF) et fantasy (FY). Cette évolution des corpus n'était pas prévue au départ mais a constitué une vraie valeur ajoutée à la réalisation de nos objectifs linguistiques et stylistiques. Une Base de données PhraseoBase a été également réalisée (cf. E.3.6.)

Des outils plus performants et ergonomiques (Lexicoscope 2.0.) ont été développés permettant une sélection fine de sous-corpus en fonction des métadonnées (genre, titre, année, auteur, traduction). Un script pour la complétion automatique des fichiers de comparaison des 6 sous-genres a été également créé. Les ALR (contenant entre 3 et 7 éléments) extraits à partir des corpus syntaxiquement annotés, ont été regroupés avec le logiciel Word2vec en *motifs textuels*.

#### 4.2. Volet linguistique

Un des résultats marquants de ce volet a été l'élaboration d'une *grille sémantique* (Goossens et al., 2020) qui a permis d'annoter les ALR extraits des corpus (plus de 11 000 expressions pour le français et l'anglais). Ainsi, par ex., nos résultats ont montré qu'il était possible de distinguer la *FY* et la *SF* (deux genres de l'imaginaire) à travers des expressions appartenant à différentes catégories sémantiques (*action, communication, cognition, mouvement*, etc). Dans la *FY* des expressions renvoyant à des *actions* violentes (*briser la nuque | les os, cut his throat*) sont statistiquement représentatives, tandis que, dans la *SF*, ce sont les expressions autour de la catégorie sémantique *cognition* qui sont les plus saillantes : *X prendre | avoir (alors | brusquement | soudain) conscience | ; to solve problem, jump to conclusion, consider the possibilities*.

Un autre résultat intéressant concerne *la comparaison des pratiques rédactionnelles dans les romans anglais et français contemporains*. Par ex., nous avons observé le suremploi par les



auteurs anglais de romans fantastiques (FY) contemporains du verbe de parole neutre *to say* (63% de tous les verbes *dicendi* extraits tel, *affirm, state, admit, agree*, comparé à la grande variété lexicale pratiquée par les auteurs français de FY (l'emploi de verbes plus expressifs comme *déplorer, balbutier, clamer, glousser, cracher* (Diwersy et al. 2020).

Enfin, parmi les motifs, nous avons identifié le motif de la cigarette (*allumer, écraser, jeter, griller* (nerveusement) une (nouvelle) cigarette contre le mur, sous le bureau) comme spécifique aux romans policiers (POL) où les personnages fument beaucoup, comparé à la littérature GEN ou SENT (Gonon, Goossens, Novakova, à paraître). Plus généralement, il a été montré (Novakova & Siepmann, 2020) que la notion de motif assure le lien entre le *micro-niveau* (les récurrences phraséologiques spécifiques) et le *macro-niveau* (celui du script narratif ou fictionnel, Baroni 2007).

#### 4.3. Volet stylistique

Les apports des stylisticiens peuvent être résumés comme suit : en amont, ils ont aidé à la constitution des corpus et à la répartition des textes en sous-genres. En aval, ils ont élaboré une grille d'analyse fine (pour le français et l'anglais) des fonctions discursives des motifs. Cette approche a permis de dépasser les approches « auteuristes » en stylistique classique et de définir « autrement » les genres romanesques à travers des séquences phraséologiques. Ainsi par ex. le motif de la fenêtre (*regarder par la fenêtre*) commun au GEN et au POL (Vidotto&Goossens, 2019) a permis de rapprocher deux sous-genres que l'on a tendance à opposer hiérarchiquement au sein du champ littéraire. Les stylisticiens membres du projet ont également entamé une riche réflexion sur les approches méthodologiques innovantes dans le domaine émergent de la *Digital Stylistics* (Jacquot, Vidotto, Gonon, 2019).

### C.5 EXPLOITATION DES RESULTATS

Les résultats du projet sont exploitables à travers la Base de Données PhraseoBase (<https://phraseorom.univ-grenoble-alpes.fr/accueil>) (Cf. E.3.6.).

### C.6 DISCUSSIONS

Les objectifs du projet ont été pour la plupart réalisés, voir dépassés. Ses résultats constituent un apport à la clarification et à l'exportation de la notion de *motif textuel*, inconnue en linguistique et stylistique anglo-saxonnes. PhraseoRom constitue aussi une contribution à l'élaboration d'une "théorie opérationnelle des genres"(Rastier 2011: 72), différente de la définition thématique à travers les univers de référence. Ses résultats pourront être utilisés aussi en narratologie cognitive, en traductologie, dans des ateliers d'écriture créative, etc. Les outils créés dans le cadre du projet auront un impact dans le domaine de l'apprentissage automatique des textes (*deep learning*) et des nouvelles technologies pour la *fouille des textes*.

### C.7 CONCLUSIONS

La complémentarité et les compétences des équipes françaises et allemandes ont constitué une vraie valeur ajoutée à la réalisation du programme scientifique et à la valorisation des résultats du projet bien au-delà des deux pays. Le caractère innovant de ce projet interdisciplinaire en Humanités numériques permettra de faire évoluer l'utilisation du numérique dans la culture, l'enseignement supérieur et l'éducation et ce, grâce à l'accélération du transfert de la recherche universitaire vers la société (retombées sociétales).

## C.8 REFERENCES

- Adam, J.-M. (1992) 2011. *Les textes : types et prototypes*. 3rd ed. Paris : Armand Colin.
- Baroni, R. 2007. *La Tension narrative: suspense, curiosité et surprise*. Paris: Seuil.
- Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Diwersy, S., L. Gonon, V. Goossens, M. Gymnich, and A. Tutin, "Speech Verbs in French and English Novels", *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*: Londres, Palgrave Mcmillan, pp. 83-113, 2020.
- Gonon L., Goossens V., Novakova I. (à paraître). « Les phraséologismes spécifiques à deux sous-genres de la paralittérature : le roman sentimental et le roman policier ». B. Buffard-Morret, S. Mejri & L. Meneses (eds) *La phraséologie française en questions ?* Collection "Vertige de la langue", Hermann, Paris.
- Goossens, V, Jacquot C, Dyka S. (2020). « Science Fiction versus Fantasy: A Semantic Categorization and its Contribution to Distinguishing Two Literary Genres, in Novakova I. & D. Siepman (eds) *Phraseology and style in subgenres of the novel : a synthesis of corpus and literary perspectives*. Palgrave Macmillan (Springer), 189-222.
- Gymnich, M, B. Neumann & A. Nünning (eds). 2007. *Gattungstheorie und Gattungsgeschichte*. Trier: WVT.
- Hoey, M; 2005. *Lexical Priming: A New Theory of Words and Language*. London, New York: Routledge.
- Jacquot, C., I. Vidotto, and L. Gonon 2019, "Digital stylistic analyses in "PhraseoRom": methodological and epistemological issues in a multidisciplinary project", *International Conference Digital Stylistics in Romance Studies and Beyond*, Würzburg, 27/02-02/03/2019.
- Kraif, O.. 2016. "Le Lexicoscope: un outil d'extraction des séquences phraséologiques basé sur des corpus arborés." *Cahiers de lexicologie* 108: 91-106.
- Kraif, O. & Diwersy, S. (2012). Le Lexicoscope : un outil pour l'étude de profils combinatoires et l'extraction de constructions lexico-syntaxiques, in *Actes de la Conférence TALN 2012* (p. 399-406), Grenoble.
- Legallois, D. 2006. "Quand le texte signale sa structure: la fonction textuelle des noms sous-spécifiés." *Corela* 5. <https://journals.openedition.org/corela/pdf/1465>
- Longrée, D, & S.Mellet. 2013. "Le Motif: une unité phraséologique englobante? Étendre le champ de la phraséologie de la langue au discours.", *Langages* 189: 68-80.
- Novakova & Siepman (eds), 2020 *Phraseology and style in subgenres of the novel : a synthesis of corpus and literary perspectives*. Palgrave Macmillan (Springer), 298 p.
- Philippe, G. & Piat J. (2009). *La langue littéraire. Une histoire de la prose en France de Gustave Flaubert à Claude Simon*. Paris: Fayard.
- Quiniou, S, P Cellier, Th Charnois & D. Legallois. 2012. "Fouille de données pour la stylistique: l'exemple des motifs émergents." In *Actes des 11es journées internationales d'analyse statistique des données textuelles*, Liège, 13-15 June 2012, 821-33.
- Rastier, F. 2011. *La Mesure et le grain: sémantique de corpus*. Paris: Honoré Champion.
- Scott, M. & Ch. Tribble. 2006. *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: Benjamins.
- Siepman, D. 2015. "A Corpus-Based Investigation into Key Words and Key Patterns in Post-war Fiction." *Functions of Language* 22, no. 3: 362-99.
- Sinclair, J. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Stubbs, M. & I. Barth. 2003. "Using Recurrent Phrases as Text-Type Discriminators: A Quantitative Method and Some Findings." *Functions of Language* 10, no. 1: 61-104.
- Vidotto I. & V. Goosses (2019). « D'une fenêtre à l'autre. Étude d'un motif spécifique à la littérature blanche et au roman sentimental », *Colloque international Phraseology and Stylistics of literary language*, Université Erlangen, 13-15 mars 2019.

## D LISTE DES LIVRABLES

Non pertinent pour ce projet

## E IMPACT DU PROJET

### E.1 INDICATEURS D'IMPACT

**Nombre de publications et de communications (à détailler en E.2)**

90 réalisations au total réparties comme suit :

		Publications multipartenaires	Publications monopartenaires
International	Revue à comité de lecture	1	7
	Ouvrages ou chapitres d'ouvrage	17	7
	Communications (conférence)	3	22 (à l'étranger)
France	Revue à comité de lecture		7
	Ouvrages ou chapitres d'ouvrage		5
	Communications (conférence)		20 (dont 12 présentées à des Colloques internationaux en France)
Actions de diffusion	Articles vulgarisation		
	Conférences vulgarisation		1
	Autres		

**Autres valorisations scientifiques (à détailler en E.3) .**

	Nombre, années et commentaires (valorisations avérées ou probables)
Brevets internationaux obtenus	
Brevet internationaux en cours d'obtention	
Brevets nationaux obtenus	
Brevet nationaux en cours d'obtention	
Licences d'exploitation (obtention / cession)	
Créations d'entreprises ou essaimage	
Nouveaux projets collaboratifs	3 (avec la Section de français de l'UNIL, Prof. G. Philippe) la COST Action 2020 (Prof. Ch. Schöch, Trier) ; l'U.M.R. Litt & Arts (UGA)
Colloques scientifiques	2 Colloques internationaux en 2019 à Erlangen et à Montpellier, 2 Journées d'études (à Cracovie 2016 et à Grenoble 2016)
Autres (préciser)	4 Ateliers d'écriture créative (2 à Grenoble et 2 à Marseille) en 2018, co-organisés avec l'Association <i>Les mots voyageurs</i>

## E.2 LISTE DES PUBLICATIONS ET COMMUNICATIONS

### I. Publications multipartenaires (Internationales) :

#### Revues :

- (1) Muryn, T., and I. Novakova, "De la phraséologie aux genres textuels : état des recherches et perspectives méthodologiques", *Revue Gerflint - Synergies Pologne*, No 14, 2017.

#### Ouvrages:

- (2) Novakova, I., and D. Siepmann (eds). *Phraseology and style in subgenres of the novel : a synthesis of corpus and literary perspectives*, London, Palgrave Mcmillan, 2020, 295 p., paru en septembre 2019
- (3) Fesenmeier, L., and I. Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, 2020, à paraître, Francfort : Peter Lang.

#### Chapitres d'ouvrages:

- (4) Diwersy, S., L. Gonon, V. Goossens, M. Gymnich, and A. Tutin, "Speech Verbs in French and English Novels", in Novakova, I., and D. Siepmann (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave Mcmillan, pp. 83-113, 2020.
- (5) Dyka, S., L. Fesenmeier, and M. Gymnich, "Dans un état de NP and in a state of NP: Bridging the Syntagmatic Gap in English and French Fiction", in Novakova, I., and D. Siepmann (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave McMillan, pp. 251-278, 2020.
- (6) Dyka, S., I. Novakova, and D. Siepmann, "A Web of Analogies: Depictive and Reaction Object Constructions in Modern English and French Fiction", *Computational and Corpus-Based Phraseology. International Conference Europhras 2017*, vol. Lecture Notes in Computer Science, vol 10596, Cham, Londres, Springer, pp. 87-101, 2017.
- (7) Gallagher, J., "The representation of smiling in postwar French and English literary prose », Fesenmeier & Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Berlin, Peter Lang, (à paraître).
- (8) Gonon, L., and O. Kraif, "French and American Science Fiction During the Nineties: A Contrastive Study of Fiction Words and Phraseology", in Novakova, I., and D. Siepmann (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave Mcmillan, pp. 151-188, 2020.
- (9) Goossens, V., C. Jacquot, and S. Dyka, "Science Fiction versus Fantasy: A Semantic Categorization and its Contribution to Distinguishing Two Literary Genres", in Novakova, I., and D. Siepmann (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave Mcmillan, pp. 189-221, 2020.
- (10) Grossmann, F., M. Gymnich, and D. Siepmann, "Alcohol and Tobacco Consumption in English and French Novels since the 1950s: A Corpus-stylistic

Analysis", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave Macmillan, pp. 114-150, 2020.

- (11) Kraif, O., I. Novakova and D. Siepman, "Multi-Word Expressions and the Construction of Meaning in French and English Crime Fiction", F. Coulouma & C. Violain (eds) *Paramétrer le sens en anglais et en français. Marqueurs et structures: articulations et constructions*, Limoges, Lambert- Lucas, pp. 23-37, 2019.
- (12) Legallois, D., and S. Koch, "The Notion of Motif Where Disciplines Intersect: Folkloristics, Narrativity, Bioinformatics, Automatic Text Processing and Linguistics", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave Mcmillan, pp. 17-46, 2020.
- (13) Muryn, T., M. Niziołek, and A. Hajok, "Que dit la main dans le roman sentimental et policier ? ", Fesenmeier, L. & I. Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Francfort Peter Lang, (à paraître)
- (14) Novakova, I., and D. Siepman, "Constructions descriptives complexes dans la littérature anglaise et française ", I. Bourov and G. Fiorentino (eds) *Complexité des structures et des systèmes linguistiques : le cas des langues romanes*, Sofia, CU Romanistika, pp. 287-300, 2019.
- (15) Novakova, I., and D. Siepman, "Literary Style, Corpus stylistic and Lexico-grammatical narrative patterns. Toward the concept of literary motifs", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave McMillan , pp. 1-15, 2020.
- (16) Novakova, I., and D. Siepman, "Towards an interdisciplinary approach in the distinction for differentiating of contemporary fiction subgenres", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave McMillan, pp. 279-285, 2020.
- (17) Novakova, I., D. Siepman, and M. Gymnich, "Key adverbs and adverbial motifs in English fiction and their French functional equivalents", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave McMillan, pp. 47-81, 2020.
- (18) Sorba, J., L. Gonon, S. Dyka, and V. Goossens, "Reading and Writing as Motifs in English and French Literary Fiction", in Novakova, I., and D. Siepman (eds) *Phraseology and style in subgenres of the novel: a synthesis of corpus and literary perspectives*, London, Palgrave McMillan, pp. 223-249, 2020.

#### Communications/ Conférences :

- (19) Dyka, S., I. Novakova, and D. Siepman, "Computational and Corpus-based Phraseology – recent advances and interdisciplinary approaches", *International Conference Europhras 2017*, 13-14 novembre 2017, Londres, 2017.
- (20) Novakova, I., and D. Siepman, "Motifs textuels à pivot adverbial dans le roman contemporain anglais et français", Colloque international LLT 2018, 25-28 septembre 2018, Université Grenoble Alpes, 2018.
- (21) Novakova, I., D. Siepman, and S. Dyka, "Constructions descriptives complexes dans la littérature anglaise et française contemporaine", Colloque international

## II. Publications monopartenaires

### Revue (internationales)

- (22) Diwersy, S., and D. Legallois, "L'apport de la méthode des motifs aux analyses phraséologiques en discours", *Theorie und Empirie in der Phraseologie: StauFFenburg Verlag*, pp. 115-132, 2019.
- (23) Gonon, L., and J. Sorba, "Phraséologismes spécifiques dans les romans historiques et les romans de littérature blanche", *Journal of French Language Studies*, pp. 1-20, 2019.
- (24) Grossmann, F., "Du rituel social aux motifs : la scène de présentation interpersonnelle dans le roman", *Revue Gerflint - Synergies Pologne*, pp. 13-28, 2017.
- (25) Kraif, O., "Traduire le polar : une étude textométrique comparée de la phraséologie du roman policier en français source et cible", *Revue Gerflint - Synergies Pologne*, n°14, pp. 43-60, 2017.
- (26) Kraif, O., and J. Sorba, "Spécificités des expressions spatiales et temporelles dans quatre sous-genres romanesques (policier, science-fiction, historique et littérature générale)", *JADT' 2018 Proceedings of the 14th International Conference on Statistical Analysis of Textual Data (12-14 juin 2018)*, vol. 1, Rome, Italy, Tor Vergata University & Sapienza University, pp. 392-399, 2018.
- (27) Novakova, I., and J. Sorba, "À la recherche de motifs autour de *meurtre, crime, assassinat* dans le roman policier contemporain", *Revue Gerflint - Synergies Pologne*, n°14, pp. 29-42, 2017.
- (28) Siepmann, D., "A corpus-based investigation into key words and key patterns in post-war fiction", *Functions of language*, vol. 22, pp. 362--399, 2015.

### Revue (France)

- (29) Delente, É., and D. Legallois, "La répétition littérale dans les Rougon-Macquart : présentation et interprétation d'un phénomène peu connu", *Excavatio*, vol. XXVIII, 2016.
- (30) Grossmann, F., "Les phraséologismes montrés par comme on dit dans les romans contemporains français", *Le Français Moderne*, pp. 83-96, 2018.
- (31) Kraif, O., "Le lexicoscope : un outil d'extraction des séquences phraséologiques basé sur des corpus arborés", *Cahiers de lexicologie*, vol. 1, pp. 91-106, 2016.
- (32) Kraif, O., I. Novakova, and J. Sorba, "Constructions lexico-syntaxiques spécifiques dans le roman policier et la science-fiction", *Lidil*, vol. 53, pp. 143-159, 2016.
- (33) Legallois, D., T. Charnois, and T. Poibeau, "Repérer les clichés dans les romans sentimentaux grâce à la méthode des "motifs", *Lidil*, vol. 53, pp. 95-117, 2016.
- (34) Siepmann, D., "Lexicologie et phraséologie du roman contemporain : quelques pistes pour le français et l'anglais", *Cahiers de lexicologie*, vol. 1, pp. 21-41, 2016.

- (35) Tutin, A., and O. Kraif, "Routines sémantico-rhétoriques dans l'écrit scientifique de sciences humaines : l'apport des arbres lexico-syntaxiques récurrents", *Lidil*, vol. 53, pp. 119–141, 2016.

### Chapitres d'ouvrages internationaux

- (36) Jacquot, C., V. Goossens, and I. Novakova, "Étude contrastive des motifs spécifiques aux romans sentimentaux et aux romans Harlequin", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Berlin, Peter Lang, e (à paraître)
- (37) Grossmann, F., "La fragmentation de l'identité dans les rituels de présentation : le cas Modiano", In Krzyzanowska, Anna; Rachwalska von Rejchwald, Jolanta (eds): *Texte, Fragmentation, Créativité I. Text, Fragmentation, Creativity I*, 32, pp. 29-40, Peter Lang,
- (38) Gymnich, M., and S. Dyka, "Falling in love vs. sich verlieben in: a corpus-based analysis of <sup>1</sup><sub>SEP</sub>English and German romance novels", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Francfort, Peter Lang, (à paraître)
- (39) Jacquot, C., V. Goossens, and I. Novakova, "Étude contrastive des motifs spécifiques aux romans sentimentaux et aux romans Harlequin", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Francfort, Peter Lang, (à paraître)
- (40) Sorba, J., "Les expressions spatiales : des unités phraséologiques pour distinguer les genres romanesques ?", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of the Literary Language / Phraséologie et stylistique de la langue littéraire*, Francfort, Peter Lang, (à paraître)
- (41) Tutin, A., and N. Gharbi, "Phrases préfabriquées à fonction expressive dans les dialogues de romans contemporains", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Francfort Peter Lang, (à paraître)
- (42) Vidotto, I., and V. Goossens, "D'une fenêtre à l'autre. Étude d'un motif spécifique à la littérature blanche et au roman sentimental", Fesenmeier&Novakova (eds) *Phraseology and Stylistics of Literary Language/Phraséologie et stylistique de la langue littéraire*, Berlin, Peter Lang.

### Chapitres d'ouvrages (France)

- (43) Gonon, L., O. Kraif, I. Novakova, J. Piat, and J. Sorba, "Sur la scène de crime : Enquête sur les enjeux linguistiques et stylistiques de motifs récurrents dans le thriller contemporain", 5<sup>ème</sup> *Congrès Mondial de Linguistique Française*, 4-8 juillet 2016, Tours, Paris, Institut de Linguistique Française, [Ouvrage en ligne] : <https://doi.org/10.1051/shsconf/20162706006>, pp. 1-14, 2016.
- (44) Gonon, L., V. Goossens, O. Kraif, I. Novakova, and J. Sorba, "Motifs textuels spécifiques au genre policier et à la littérature « blanche »", 6<sup>e</sup> *Congrès Mondial de*

*Linguistique Française*, 9-13 juillet 2018, Mons, Paris, Institut de Linguistique Française, [Ouvrage en ligne] : <https://doi.org/10.1051/shsconf/20184606007>, pp. 1-14, 2018.

- (45) Gonon, L., V. Goossens, and I. Novakova, "Les phraséologismes spécifiques à deux sous-genres de la paralittérature : le roman sentimental et le roman policier", *La phraséologie française en questions ?*, vol. Vertige de la langue, Paris, Hermann, (à paraître)
- (46) Goossens, V., D. Legallois, and I. Novakova, "Aperçu de quelques méthodes pour l'extraction des constructions en corpus", J. François (ed) *L'expansion pluridisciplinaire des grammaires de constructions*, Caen, Presses universitaires de Caen (à paraître).
- (47) Legallois, D., "Caractériser le style d'un auteur par des patrons lexicogrammaticaux : une nouvelle approche en stylistique", M. Monte, Ph. Wahl et S. Thonnerieux (eds) *Stylistique et méthode. Quels paliers de pertinence textuelle ?*, Presses universitaires de Lyon, collection Textes et langue, pp. 87-98.

### Communications/conférences (Internationales)

- (48) Diwersy, S., and O. Kraif, "Patterns and Novels – an outline of the PhraseoRom project and its provisional results", Colloque international *Digital Stylistics in Romance Studies and Beyond*, Würzburg, 27/02-02/03/2019.
- (49) Diwersy, S., and O. Kraif, "Motifs et romans dans le projet PhraseoRom – jalons méthodologiques et perspectives de résultats", Colloque international *Phraseology and Stylistics of Literary Language*, Erlangen, 13-15/03/2019.
- (50) Gonon, L., and J. Sorba, "Phraséologismes autour du vin et du champagne dans le roman contemporain", Colloque international *Terminologies et discours gastronomiques et œnologiques. Le vin et les autres produits*, Faculté de philologie de l'Université d'Innsbruck (Autriche), 19-20/09/2019.
- (51) Goossens, V., C. Jacquot, and I. Novakova, "Motifs textuels spécifiques aux romans sentimentaux et Harlequin (en contraste)", Colloque international *Phraseology and Stylistics of Literary Language*, Erlangen, Germany, 13-15/03/2019.
- (52) Grossmann, F., "« Se présenter » et « présenter quelqu'un à autrui » : formes et fonctions des motifs phraséologiques de l'auto-présentation et de la présentation d'autrui dans le roman : l'exemple de Simenon et de Modiano", *Journée phraséologique franco-polonaise - 18 octobre 2016*, Université pédagogique de Cracovie, 10/2016.
- (53) Grossmann, F., "La fragmentation de l'identité dans les rituels de présentation : le cas Modiano.", Colloque international *Texte, Fragmentation, Créativité*, Lublin (Pologne), 21-22 octobre 2016, 2016.
- (54) Jacquot, C., I. Vidotto, and L. Gonon, "Digital stylistic analyses in "PhraseoRom": methodological and epistemological issues in a multidisciplinary project", International Conference *Digital Stylistics in Romance Studies and Beyond*, Würzburg, 27/02-02/03/2019.
- (55) Jacquot, C., and V. Goossens, "Stylistic boundaries between French science-fiction and fantasy novels", International Conference *PALA "Stylistics without Borders"*, Liverpool, Angleterre, 10-14/07/2019.



- (56) Kraif, O., "Traduire le polar : une étude textométrique comparée de la phraséologie du roman policier en français source et cible", *Journée phraséologique franco-polonaise*, 18 octobre 2016, Université pédagogique de Cracovie, 2016.
- (57) Kraif, O., and J. Sorba, "Spécificités des expressions spatiales et temporelles dans quatre sous-genres romanesques (policier, science-fiction, historique et littérature générale)", *JADT' 2018- 14th International Conference on Statistical Analysis of Textual Data*, Rome, Italy, 12-14 juin 2018.
- (58) Legallois, D., "« Je vais partir pour Rome et je n'ai pas, malheureusement, le temps de répondre à vos questions » – la traduction des répétitions dans *Les Rougon-Macquart*", Colloque international « *Traduire Zola, du XIXe siècle à nos jours* », 27-29 avril 2017, Università Roma Tre, 2017.
- (59) Legallois, D., and S. Diwersy, "Analyzing Idiolectal Evolution from Literary Corpora", Colloque international *Phraseology and Stylistics of Literary Language*, Erlangen, 13-15/03/2019.
- (60) Novakova, I., "Constructions lexico-syntaxiques spécifiques dans le roman policier et la science-fiction", Conférence invitée à la Section de français de l'Université de Lausanne, dirigée par le Prof. Gilles Philippe, Lausanne, 11-05-2016.
- (61) Novakova, I., La phraséologie spécifique du roman contemporain français, Conférence invitée au Centre d'études francophones, Université de Sofia, Bulgarie, 07/06/2017.
- (62) Novakova, I., A la recherche de motifs textuels pour l'identification des genres textuels, Conférence invitée à l'Académie des Science (section de Phraséologie), Varsovie, Pologne, 05/04/2019.
- (63) Novakova, I., Phraséologie et stylistique des genres littéraires. Le projet PhraseoRom, , Conférence invitée au Département de Français de l'Institut de Communication et de Culture de l'Université d'Aarhus, Danemark, 30-10-2019.
- (64) Novakova, I., and J. Sorba, "À la recherche de motifs autour de meurtre, crime, assassinat dans le roman policier contemporain", *Journée phraséologique franco-polonaise*, 18 octobre 2016, Université pédagogique de Cracovie, 2016.
- (65) Siepmann, D., "Tolkien as a Stylist: a Corpus-Based Investigation into 'The Lord of the Rings'", *9th International Corpus Linguistic Conference*, 24-28 juillet 2017, University of Birmingham, 2017.
- (66) Siepmann, D., and S. Dyka, "A web of analogies: key descriptive constructions in modern fiction", *9th International Corpus Linguistic Conference*, 24-28 juillet 2017, University of Birmingham, 2017.
- (67) Sorba, J., "Les expressions spatiales : des unités phraséologiques pour distinguer les genres romanesques ?", Colloque international *Phraseology and Stylistics of Literary Language*, Erlangen, Germany, 13-15/03/2019.
- (68) Tutin, A., and N. Gharbi, "Les phrases préfabriquées à fonction expressive dans les dialogues de romans contemporains", Colloque international *Phraseology and Stylistics of Literary Language*, Erlangen, 13-15/03/2019.
- (69) Vidotto, I., and V. Goossens, "D'une fenêtre à l'autre. Étude d'un motif spécifique à la littérature blanche et au roman sentimental", Colloque international *Phraséologie et stylistique de la langue littéraire*, Erlangen, Germany, 13-15/03/2019.

## Communications /Conférences Colloques en France

- (70) Chambre, J., and O. Kraif, "Séquences récurrentes typiques du roman policier et de science-fiction : une étude préliminaire sur corpus", *Journées internationales de la linguistique de corpus*, 4-6 juillet 2017, Université Grenoble Alpes, 2017.
- (71) Diwersy, S., "ALR, extensions collocationnelles, motifs : enjeux méthodologiques et perspectives du projet PhraseoRom", Colloque international *Interroger le texte à l'ère de l'"intelligence mécanique" : la stylistique outillée, au carrefour du disciplinaire et de l'interdisciplinaire*, Montpellier, France, 11-12/06/2019.
- (72) Diwersy, S., "Présentation de la plateforme PhraseoRom", Colloque international *Interroger le texte à l'ère de l'"intelligence mécanique" : la stylistique outillée, au carrefour du disciplinaire et de l'interdisciplinaire*, Montpellier, France, 11-12/06/2019.
- (73) Gonon, L., O. Kraif, I. Novakova, J. Piat, and J. Sorba, "Sur la scène de crime : Enquête sur les enjeux linguistiques et stylistiques de motifs récurrents dans le thriller contemporain", *5<sup>ème</sup> Congrès Mondial de Linguistique Française*, 4-8 juillet 2016, Tours, Paris, Institut de Linguistique Française,
- (74) Gonon, L., V. Goossens, and I. Novakova, "Les phraséologismes spécifiques à deux sous-genres de la paralittérature : le roman sentimental et le roman policier", Colloque international *Phraséologie française*, 21-22 septembre 2017, Université d'Artois, Arras, 2017.
- (75) Gonon, L., V. Goossens, O. Kraif, I. Novakova, and J. Sorba, "Motifs textuels spécifiques au genre policier et à la littérature « blanche »", *6e Congrès Mondial de Linguistique Française*, 9-13 juillet 2018, Mons, Paris, Institut de Linguistique Française,
- (76) Gonon, L., and J. Sorba, "Phraséologismes spécifiques dans les romans historiques et les romans de littérature blanche", Conférence internationale LLT 2018, 25-28 septembre 2018, Université Grenoble Alpes, 2018.
- (77) Gonon L. and I. Vidotto « Corpus romanesques parallèles : études sur les corpus anglais et français du projet PhraseoRom », Journée d'étude *Humanités numériques et texte littéraire traduit. Editer et analyser des corpus de versions parallèles*, Grenoble, le 24 septembre 2019.
- (78) Goossens, V., C. Jacquot, and I. Novakova, "Distinguer les genres de la littérature de l'imaginaire par les motifs textuels", Journée d'étude *"Lexique et frontières de genres"*, Pau, France, 2018.
- (79) Goossens, V., D. Legallois, and I. Novakova, "Building Fictional Worlds with Linguistic Constructions", *Tenth International Conference on Construction Grammar*, Paris, 16-18/07/2018.
- (80) Grossmann, F., "Les phraséologismes montrés dans les romans contemporains", *Journée de valorisation Agir Phraséotext*, Grenoble, 8 décembre 2016, 2016.
- (81) Grossmann, F., "Contraster les styles à l'aide d'une approche outillée : les motifs de la lecture et de l'écriture chez Modiano et Houellebecq", Colloque international *Interroger le texte à l'ère de l'"intelligence mécanique" : la stylistique outillée, au carrefour du disciplinaire et de l'interdisciplinaire*, Montpellier, France, 11-12/06/2019.
- (82) Jacquot, C., and I. Vidotto, "Enjeux méthodologiques et épistémologiques d'une annotation stylistique : l'exemple du projet PhraseoRom", Colloque international *Interroger le texte à l'ère de l'"intelligence mécanique" : la stylistique outillée, au carrefour du disciplinaire et de l'interdisciplinaire*, Montpellier, France, 11-12/06/2019.

- (83) Kraif, O., "Autour de l'extraction d'arbres lexico-syntactiques récurrents spécifiques : corpus, outils et méthodes", *Journée de valorisation Agir Phraséotext*, Grenoble, 8 décembre 2016, 2016.
- (84) Kraif, O., "Sous-genre littéraire et traduction : une approche textométrique", *Congrès mondial de traductologie, Atelier textométrie et traduction*, 17 avril 2017, Université Paris Ouest – Nanterre – La Défense, 2017.
- (85) Kraif, O., and I. Novakova, "Analyse linguistique du lexique sur grands corpus informatisés", Cycle de conférences "Les méthodes évoluent, et vous?", Conférence invitée à la MSH Alpes, Grenoble, 14-06-2016. (vulgarisation)
- (86) Legallois, D., "Pour en revenir au texte: aspects de la lecture à « bonne distance » en humanités numériques littéraires », Colloque « *Des humanités numériques littéraires ?* », 15-22 juin 2017, Cerisy, 2017.
- (87) Legallois, D., "La discrimination des genres textuels: du lexical au grammatical", Keynote talk, Colloque international *Lexique et frontières de genre*, Bayonne, 10/10/2019.
- (88) Novakova, I., "Le projet PhraseoRom : à la croisée de la phraséologie et des genres littéraires", Keynote talk au Colloque international *Interroger le texte à l'ère de l'"intelligence mécanique" : la stylistique outillée, au carrefour du disciplinaire et de l'interdisciplinaire*, Montpellier, France, 11-12/06/2019.
- (89) Novakova, I., "Identifier et définir les genres littéraires à travers la phraséologie étendue", Colloque international *Lexique et frontières de genres*, Bayonne, 10/10/2019.
- (90) Piat, J., and L. Gonon, "Analyse stylistique de phraséologismes spécifiques dans le roman contemporain : réflexions méthodologiques", *Journée de valorisation Agir Phraséotext*, jeudi 8 décembre 2016, 2016.

### E.3 LISTE DES ELEMENTS DE VALORISATION

Dans le but de faire connaître et valoriser les résultats du projet PhraseoRom, plusieurs actions et travaux ont été réalisés entre 2016 et 2019 (consultables sur le site *trilingue* du projet <https://phraseorom.univ-grenoble-alpes.fr/>), qui peuvent être regroupés comme suit :

#### E.3.1. Publications

Les membres des équipes françaises et allemandes ont réalisé au total 90 travaux (individuels et collectifs), dont 2 ouvrages internationaux.

L'ouvrage *Phraseology and Style in Subgenres of the Novel. A Synthesis of Corpus and Literary Perspectives* (Novakova&Siepmann eds 2020), paru en anglais chez Palgrave MacMillan en septembre 2019, n'était pas prévu dans notre programme de départ. Il a permis de valoriser une partie de nos résultats dans la communauté des chercheurs en linguistique et stylistique francophones et anglophones. <https://www.palgrave.com/gp/book/9783030237431>

Un deuxième ouvrage *Phraseology and Stylistics of Literary Language* (Fesenmeier & Novakova eds) paraîtra en 2020 chez Peter Lang. Il est issu d'une sélection de contributions présentées au Colloque International organisé par les membres du projet les 13-15 mars 2019 à la Friedrich Alexander Universität (FAU) Erlangen-Nuremberg, Allemagne). Ce livre confrontera différentes approches de la linguistique et de la stylistique pour l'analyse du langage littéraire.

### **E.3.2. Conférences invitées, thèses de Doctorat et mémoires de Master**

Le projet a été également valorisé à travers une série de conférences en France et à l'étranger afin de faire connaître ses résultats à la communauté scientifique (cf. Liste des publications du No 60 à 63 et le No 85). On pourrait citer aussi la réalisation en cours de 3 thèses de Doctorat, de 6 Mémoires Master soutenus sur les thématiques de PhraseoRom (impact scientifique).

### **E.3.3. Organisation d'une Journée d'étude à Cracovie :**

Une Journée phraséologique franco-polonaise a eu lieu le 18 octobre 2016 à l'Université pédagogique de Cracovie, co-organisée par les collaborateurs extérieurs au projet (T. Muryn et M. Niziolek) et des membres de l'équipe française, à laquelle ont été présentés les premiers résultats des études pilotes, réalisées dans le cadre de PhraseoRom. Cette Journée a donné lieu à la publication du No 14 de la revue *Synergies Pologne* (T. Muryn & I. Novakova eds, 2017), intitulé *De la phraséologie aux genres textuels : état des recherches et perspectives méthodologiques*.

### **E.3.4. Organisation de deux Colloques internationaux**

Le Colloque international *Phraseology and Stylistics of Literary Language*, organisé par L. Fesenmeier et S. Koch (mars 2019, FAU, Erlangen), a réuni 45 chercheurs de 15 pays. <https://www.romanistik.phil.fau.de/tagung-phraseorom-call-for-papers-dt/>. Ce Colloque a été le lieu d'échange entre les chercheurs dans deux sections scientifiques : Section 1 : *Recherches statistiques appliquées aux grands corpus littéraires* ; Section 2 : *Analyses linguistiques et stylistiques des genres littéraires, stylistique empirique des corpus*. Conférenciers invités : Michaela Mahlberg (University of Birmingham), Dan McIntyre (University of Huddersfield) Véronique Montémont (Université de Lorraine, ATILF, France).

Un deuxième Colloque international a été également organisé à l'initiative de deux membres du projet C. Jacquot et S. Diwersy (U. Paul Valéry, Montpellier), intitulé *Interroger le texte à l'ère de l'"intelligence mécanique": la stylistique outillée au carrefour du disciplinaire et de l'interdisciplinaire* <https://www.fabula.org/actualites/colloque-interroger-le-texte-l-ere-de-l-intelligence-mecanique-la-stylistique-outillee-au-91485.php>. Ce Colloque, non prévu dans nos objectifs initiaux, a permis de lancer une riche et utile discussion sur les méthodes et les outils en Stylistique numérique (*Digital Stylistics*), discipline émergente. Il a réuni une vingtaine de chercheurs de 7 pays. Il a permis de confronter les méthodologies développées dans le cadre de PhraseoRom à celles appliquées par des collègues travaillant dans le même domaine (par ex. dans le cadre de COST Action Horizon 2020 *Distant Reading for European Literary History*, dirigée par le Ch. Schöch, Trier Universität). Conférenciers invités : Christof Schöch (Trier Universität) ; Berenike Herrmann (Basel Universität), Iva Novakova (Lidilem, UGA).

### **E.3.5. Journée de valorisation du projet PhraseoRom (19.12.2019, Université Grenoble Alpes(UGA))**

La Journée de valorisation du projet et d'inauguration de la Base de donnée PhraseoBase a réuni les membres du projet et des invités extérieurs (Th. Mettraux de l' U. de Lausanne, D. Mayaffre de l'U. de Nice). Pour plus de détails : <https://phraseorom.univ-grenoble-alpes.fr/fr/actualites/journee-valorisation-projet-et-dinauguration-base-donnees-phraseobase>

Les enregistrements des présentations et de la Table ronde sont disponibles sur <https://phraseorom.univ-grenoble-alpes.fr/fr/actualites/videos-journee-valorisation-projet-phraseorom>

### **E.3.6. La création d'une plate-forme : la Base de données (BDD) *PhraseoBase***

Cette BDD, créée par O. Kraif (UGA) et S. Diwersy (U. P. Valéry, Montpellier 3) et par les collaborateurs du projet a été mise à la disposition de la communauté de chercheurs, d'enseignants, d'étudiants et de doctorants en linguistique, en stylistique et en TAL (inauguration le 19.12.2019).

<http://phraseotext.univ-grenoble-alpes.fr/phraseobase/index.html>

Une première application donne accès, grâce au Lexicoscope, aux corpus trilingues (comparables et parallèles) [http://phraseotext.univ-grenoble-alpes.fr/lexicoscope\\_2.0/](http://phraseotext.univ-grenoble-alpes.fr/lexicoscope_2.0/)

La seconde application donne accès à un ensemble d'analyses linguistiques et stylistiques menées sur les données extraites de ces corpus. <http://phraseotext.univ-grenoble-alpes.fr/phraseoLing/#> et des visualisations des résultats (AFC, diagrammes, échelonnements multidimensionnels, etc).

Enfin, la *PhraseoBase* propose un ensemble de ressources à télécharger <http://phraseotext.univ-grenoble-alpes.fr/phraseobase/index.html>: manuels d'annotations sémantiques et stylistiques en français et en anglais, liste des étiquettes morpho-syntaxiques associées aux tokens des corpus, etc.

### **E.3.7. Développement de nouveaux partenariats,**

En parallèle du projet *PhraseoRom* a été réalisé un autre projet intitulé *PhraséoLit* (responsable scientifique I. Novakova), lauréat de l'AAP SFR Création UGA 2018 <http://maisondelacreation.univ-grenoble-alpes.fr/fr/recherche/projets/phraseolit-phraseologie-specifique-a-la-litterature-contemporaine-311423.kjsp?RH=MCFR RECH02>

*PhraséoLit* a permis de créer une collaboration originale avec L'Association de création littéraire *Les Mots Voyageurs* (Marseille) <http://lesmotsvoyageurs.com/>, dirigée par Noëlle Mathis. En utilisant une partie des résultats de *PhraseoRom*, quatre Ateliers d'écriture créative (2 à Grenoble et 2 à Marseille) réunissant 20 participants ont été organisés. Ils ont suscité un vif intérêt parmi les participants (étudiants, doctorants, enseignants, retraités). Les motifs textuels extraits des corpus *Phraseorom* (littérature générale) ont servi de déclencheurs d'écriture créative (par ex. les motifs *entrer dans un lieu* (pousser la porte)→ *entrer dans l'écriture*). Les textes recueillis des écrivains seront analysés sur le plan linguistique et stylistique. Nous souhaitons étendre l'expérience acquise lors de ces ateliers d'écriture *PhraséoLit* / *PhraseoRom* à de nouveaux ateliers, en collaboration avec la SFR Création (UGA), l'UMR Litt&Arts (UGA), l'Association *Les Mots voyageurs* (impact sociétal).

### **E.3.8. Collaborations internationales**

La diffusion des travaux réalisés dans le cadre de *PhraseoRom* a donné lieu à de multiples contacts et partenariats internationaux, qui contribuent à élargir et à consolider le réseau scientifique tissé autour du projet. Citons parmi ces contacts :

- a) L'équipe du Prof. Michaela Mahlberg (Department of English Language and Linguistics, U. de Birmingham, GB) qui a été *keynote speaker* et membre du Comité

scientifique du Colloque international d'Erlangen. Elle publiera un chapitre dans l'ouvrage *Phraseology and Stylistics of Literary Language* (Fesenmeier & Novakova eds, 2020).

- b) l'équipe COST Action Horizon 2020 *Distant Reading for European Literary History*, dirigée par le Prof. C. Schöch, Trier Universität). C. Schöch a été membre du Comité scientifique du Colloque *Phraseology and Stylistics of Literary Language* (Erlangen, mars 2019) et *keynote speaker* au Colloque international de Montpellier (juin 2019): *Interroger le texte à l'ère de l'"intelligence mécanique": la stylistique outillée au carrefour du disciplinaire et de l'interdisciplinaire*;
- c) Les collègues de l'Université pédagogique de Cracovie (Pologne) qui ont été collaboratrices extérieures du projet (organisation de la Journée d'études en 2016 à Cracovie, la publication de la revue *Synergie* 14, 2017, la publication d'un chapitre dans l'ouvrage *Phraseology and Stylistics of Literary Language* (Fesenmeier & Novakova eds, 2020). Prof. Teresa Muryn et Malgorzata Niziolek ont été membres du CS du Colloque international d'Erlangen.
- d) L'équipe du Prof. Anna Krzyżanowska de l'Institut de Philologie Romane *Marie Curie-Skłodowska*, Lublin (Pologne) dans le cadre du projet Polonium 2018-2019 sur les « Pragmatèmes en contraste: de la modélisation linguistique au codage lexicographique ». Coordinateurs : Francis Grossmann (UGA, Lidilem) et Anna Krzyżanowska (Lublin). Une partie des corpus PhraseoRom ont été utilisés, entre autres, dans ce projet;
- e) Des collègues du Département d'études romanes de l'Université de Sofia (Bulgarie), et, en particulier, avec Mme Margarita Rouski (MCF) qui utilise, dans ses derniers travaux, les corpus et la méthodologie PhraseoRom ;
- f) Des collègues de l'équipe du Prof. Gilles Philippe de la Section de français de la Faculté des lettres de l'Université de Lausanne (échange de conférences, participation de Thibaud Mettraux au Colloque d'Erlangen en mars 2019 et à la Table ronde de la Journée PhraseoRom du 19.12.2019). Ilaria Vidotto, actuellement première assistante à la Section de français de l'UNIL, a effectué un Post-doc en stylistique à l'UMR Litt&Arts, UGA, pendant 1 an (2018-2019) , sous la direction de Laetitia Gonon dans le cadre du projet PhraseoRom. Nous envisageons de soumettre un projet dans la continuité de PhraseoRom, lié aux approches numériques des textes littéraires.

#### E.4 BILAN ET SUIVI DES PERSONNELS RECRUTES EN CDD (HORS STAGIAIRES)

Identification				Avant le recrutement sur le projet			Recrutement sur le projet				Après le projet				
Nom et prénom	Sexe H/F	Adresse email (1)	Date des dernières nouvelles	Dernier diplôme obtenu au moment du recrutement	Lieu d'études (France, UE, hors UE)	Expérience prof. Antérieure, y compris post-docs (ans)	Partenaire ayant embauché la personne	Poste dans le projet (2)	Durée missions (mois) (3)	Date de fin de mission sur le projet	Devenir professionnel (4)	Type d'employeur (5)	Type d'emploi (6)	Lien au projet ANR (7)	Valorisation expérience
Johan DIDIER	H		19.12.2019	Master Industries de la langue à l'UGA	France	non	Lidilem UGA	Ingénieur d'étude	24 mois	01.10.2018	recherche d'emploi				
Vannina GOOSSENS	F		31.01.2020	Doctorat en Sciences du Langage	France	5 ans CDD	Lidilem UGA	Post-doc	34 mois	31.12..2019	CDD	enseignement et recherche publique, UGA	Post-doc sur autre projet ANR	Non	Oui
Ilaria VIDOTTO	F		31.01.2020	Doctorat en littérature	France	non	Litt&Arts UGA	Post-doc	12 mois	31.08.2019	CDD	Assistante à l'UNIL (Lausanne, Suisse)	enseignant-chercheur	oui	oui