

## Unités polylexicales et construction du sens dans le roman policier français et anglais

Olivier KRAIF (U. Grenoble Alpes– LIDILEM)

Iva NOVAKOVA (U. Grenoble Alpes– LIDILEM)

Dirk SIEPMANN (U. d'Osnabrück, Institut d'Études Anglophones)

[olivier.kraif@u-grenoble3.fr](mailto:olivier.kraif@u-grenoble3.fr)  
[iva.novakova@u-grenoble3.fr](mailto:iva.novakova@u-grenoble3.fr)  
[dirk.siepmann@uos.de](mailto:dirk.siepmann@uos.de)

# Hypothèses

- Le langage littéraire se caractérise par la **surreprésentation** de lexèmes, collocations, patrons valenciels et colligations qui lui sont spécifiques (Siepmann 2015):
  - Il en était là de ses réflexions quand...*
  - PRONOM + en + être + là + de + PRONOM POSS. + NOM désignant un processus cognitif + quand/lorsque
  
- Certaines expressions polylexicales (EP) récurrentes permettent d'identifier des structures lexico-grammaticales prototypiques d'un certain (sous-)genre textuel. Ces structures peuvent, dans certains cas, s'apparenter à des **motifs de mots co-occurents**:
  - *Un ensemble d'éléments fixes et variables susceptibles d'accompagner la structuration textuelle, et simultanément, de caractériser des textes de genres divers (Longrée & Mellet 2013 : 66)*
  
  - *Une unité multidimensionnelle, constituée à la fois d'associations lexicales et grammaticales, d'appariements entre forme et sens ou fonction pragmatique/discursive (Legallois 2012 : 45)*

# Fondements théoriques

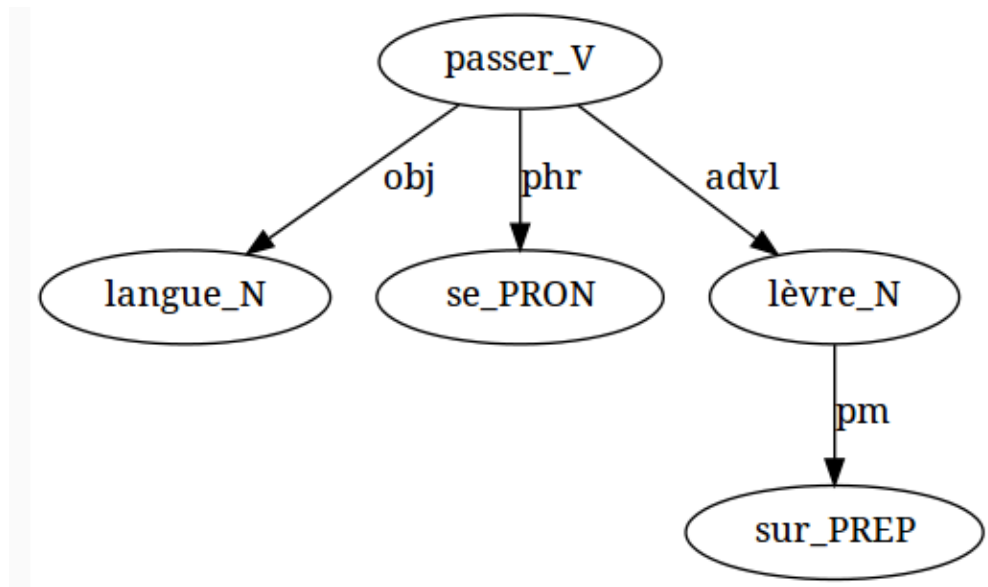
- Le motif: une construction lexico-grammaticale récurrente qui a la fonction de **marqueur discursif structurant** (Longrée & Mellet, 2012: 718)
- Modèles fonctionnels et contextualistes (Sinclair 1991, Biber 2009)
- 4 niveaux d'analyse des unités linguistiques: lexical, sémantique, syntaxique et discursif dans la perspective de l'organisation du texte littéraire (Sinclair 2004)
- Approche *corpus-driven*: observations guidées par les données
  - les outils numériques d'exploration du corpus permettent de faire apparaître des régularités et des phénomènes attendus mais parfois aussi inattendus à partir desquelles on peut définir ou caractériser les genres littéraires
  - Bakhtine (1984): la langue ne s'apprend et ne se pratique qu'à travers les genres de discours.

# Corpus

CORPUS	FRANÇAIS	ANGLAIS
LITTÉRAIRE (en nombre de mots)	16 M.	12 M.
POL	4 M.	6 M.
Nombres d'EP	3185	4777
TEXTES	35	55
AUTEURS	Chattam, Grangé, Vargas, Simenon, Aubert, Bannel, Dard	Cornwell, Bagley, Church, Christie, Martin, Will, Bayer

# Méthodologie

- Outils : Lexicoscope (Kraif & Diwersy, 2012)
- Extraction automatique des expressions polylexicales (**Arbres lexico-syntaxiques récurrents, ou ALR**) spécifiques au sous-corpus POL
- P. ex. en partant de *langue\_N* :



# Méthodologie

- Répartition de l'expression *se passer la langue sur les lèvres* :  
**POL : 12 occ.**, HIST : 2 occ., SF : 1 occ., AUTRE : 1 occ., SENT 0 occ.

Il semble qu'elle soit spécifique à POL. Comment le mesurer ?

- Calcul de *spécificité* des EP : fréquences relatives dans un sous-corpus par rapport aux fréquences dans l'ensemble du corpus. Tableau de contingence:
  - f1 : la fréquence dans POL
  - f2 : la fréquence dans tout le corpus LIT\_MOD
  - T1 : le nombre total de mots de POL
  - T2 : le nombre total de mots de LIT\_MOD

-> Calcul du *loglikelihood ratio* (rapport de vraisemblance)

-> Seuil de spécificité  $\geq 10.83$ , dispersion  $\geq 3$
- Extraction d'ALR *émergents* (cf. motifs séquentiels émergents, Quiniou et al. 2012)

# Méthodologie

Sous-arbres	f1	f2	Disp	Loglike
<l=crime,c=N,#1>&&<l=scène,c=N,#2>&&<l=de,c=PREP,#3>::(mod,2,1) (pm,2,3)	161	165	3	408,05
<l=un,c=DET,#1>&&<l=regard,c=N,#2>&&<l=lancer,c=V,#3>::(det,2,1) (obj,3,2)	121	158	6	182,75
<l=un,c=DET,#1>&&<l=coup,c=N,#2>&&<l=jeter,c=V,#3>::(det,2,1) (obj,3,2)	207	371	6	156,13
<l=champ,c=N,#1>&&<l=vision,c=N,#2>&&<l=de,c=PREP,#3>&&<l=(?:son  leur mon ton notre votre),c=PRON,#4>::(det,1,4) (pm,1,3) (mod,1,2)	23	7	4	118,06
<l=mot,c=N,#1>&&<l=sans,c=PREP,#2>&&<l=(?:un une des d de),c=DET,# 3>::(det,1,3) (pm,1,2)	142	248	6	114,15
<l=jeune,c=A,#1>&&<l=inspecteur,c=N,#2>::(attr,2,1)	42	47	4	86,72
<l=quai,c=N,#1>&&<l=orfèvre,c=N,#2>::(mod,1,2)	38	47	5	63,98
<l=tasse,c=N,#1>&&<l=café,c=N,#2>::(mod,1,2)	65	112	5	53,98
<l=crâne,c=N,#1>&&<l=son,c=PRON,#2>::(det,1,2)	135	304	6	52,45
<l=vitesse,c=N,#1>&&<l=à,c=PREP,#2>&&<l=tout,c=PRON,#3>::(det,1,3) (pm,1,2)	78	151	6	48,1
<l=meurtre,c=N,#1>&&<l=série,c=N,#2>::(mod,2,1)	22	24	3	47,98
<l=cadrer,c=V,#1>&&<l=pas,c=ADV,#2>&&<l=ne,c=ADV,#3>::(neg,1,2) (ad,2,3)	25	30	3	44,72
<l=mise,c=N,#1>&&<l=scène,c=N,#2>::(mod,1,2)	70	135	6	43,61
<l=chance,c=N,#1>&&<l=tomber,c=V,#2>&&<l=de,c=PREP,#3>::(pm,1,3) (mod,1,2)	10	5	3	41,4
<l=point,c=N,#1>&&<l=suspension,c=N,#2>&&<l=de,c=PREP,#3>::(pm,1,3) ) (mod,1,2)	22	30	3	30,44
<l=instant,c=N,#1>&&<l=à,c=PREP,#2>&&<l=même,c=A,#3>::(attr,1,3) (pm,1,2)	42	77	5	29,91

# Résultats

Pivots récurrents spécifiques (*key lexemes, key compounds, key patterns*)

Sous-classes	FRANÇAIS	ANGLAIS
Personnages	<i>lieutenant, inspecteur de police, légiste, suspect, chef de police, cadavre, victime</i>	<i>inspector, cop, public prosecutor, victim, suspect, body, superintendent</i>
Lieux	<i>commissariat, quai des Orfèvres, laboratoire de police, bureau, morgue</i>	<i>autopsy suite, office, morgue</i>
Objets/traces	<i>empreintes, traces</i>	<i>(get) fingerprints, traces</i>
Actes	<i>crime, meurtre, ...</i>	<i>crime, murder, ...</i>
Localisation temporelle	<i>Pleine nuit Minuit Fin d'après midi</i>	<i>time of the death</i>



# Scène de crime

## (récurrence thématique)

- **Collocation** la plus spécifique pour le POL fr ([loglike 408](#)) : « affinité » de *crime* pour le pivot *scène*; en revanche *scène du meurtre* : 3 occ uniquement, « aversion » du pivot *meurtre* pour le collocatif *scène* (cf. Hoey, 2005)

Sous-corpus	Occurrences
POL	116
AUT	2
SF	1

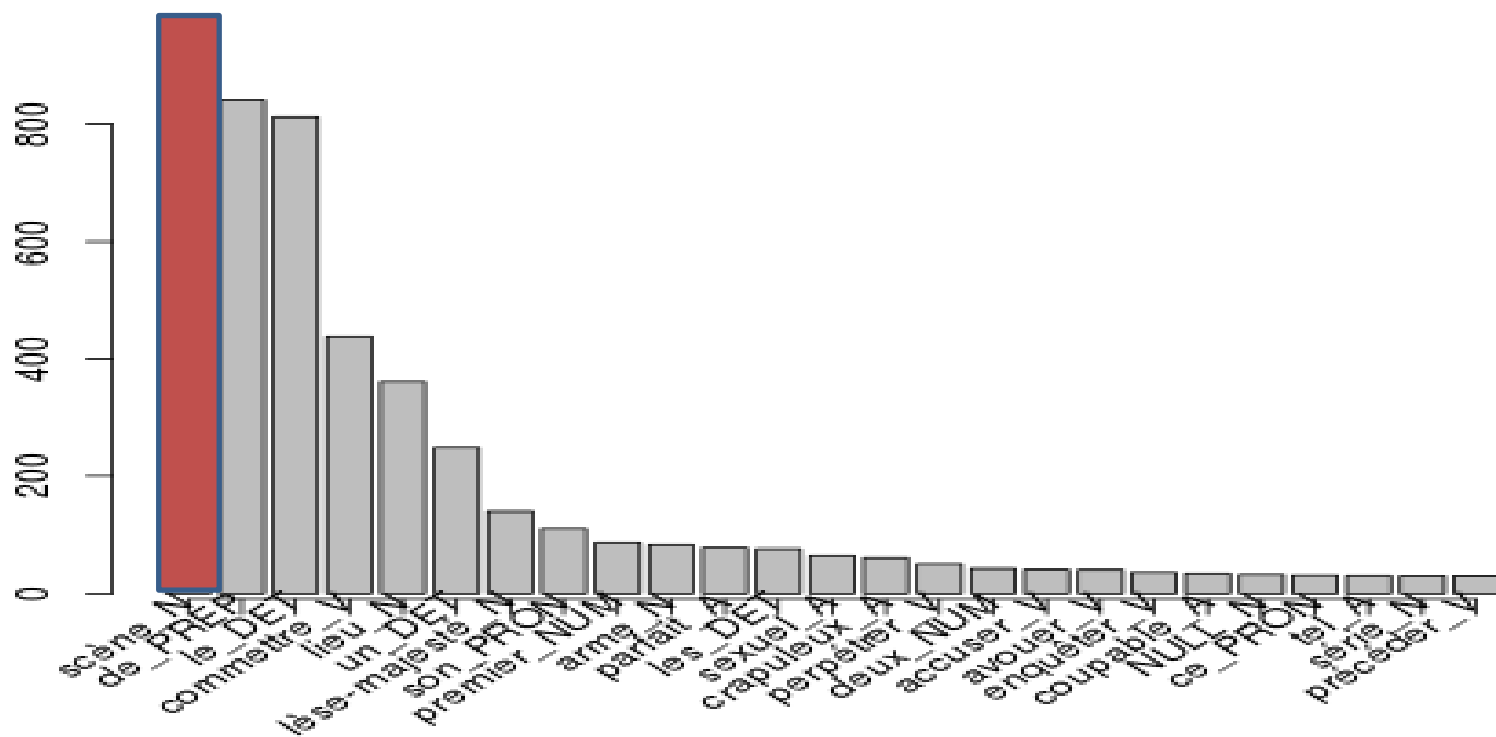
(1) **Sur la scène de crime**, Langleber me fait signe de venir. (Grangé, POL)

# Scène de crime

(récurrence thématique)

(*loglikelihood* 1155)

Lexicogramme relatif au pivot  
crime\_N



# Scene of the crime

Statistiquement moins spécifique en anglais (avec quelques variations)

Scene of the crime	Indice de spécificité <i>loglikelihood</i>
<i>the scene of the crime</i>	32
<i>a / the crime scene</i>	30
<i>on the scene</i>	20
<i>the scene of the accident</i>	6,1
<i>the scene of the death</i>	4,8

(2) I can only suppose it's because he saw me **at the crime scenes**.

(Cornwell)

(3) Driven **to the scene of the crime**, Dick Milton was furious when he caught sight of a French television crew. (The Mammoth Book)

# Scène de crime/scene of the crime

- **Colligation** la plus spécifique pour *scène de crime* (log 179) :

Prép. *sur* + Dét. (*la, les*) + *scène de crime* (SC);

*at the scene of the crime*

- **Motif de la localisation** :

- Extension **syntagmatique** du motif à gauche (préférence pour les Verbes de **mouvement** → **lieu**) :

*Vmouv* (*arriver*) + Prép (*sur*) + Dét (*la*) + *scène de crime*

(4) C'était la première chose à faire **en arrivant sur une scène de crime**.  
(Chattam, POL)

(5) By the time they **reached the crime scene**, reporters had gathered in the night, all poised as Brazil's peers tried to get close to a terrible tragedy. (Cornwell)

# Scène de crime

- Variations **paradigmatiques** du motif de la localisation :
  - Variations sur la PREP : *sur, autour de, (tout) près de, au seuil de, sur le seuil de, dans les environs de, à proximité de* + Dét. + SC

(6) Du temps où il était inspecteur, Brodin demandait régulièrement qu'on filme la foule présente **autour d'une scène de crime**. (Chattam, POL)

-> moins de variation en anglais: *on, at, near*

- Variations sur le Vmouv : *aller, se rapprocher, partir, arriver, revenir*

(7) Il attrapa sa veste et **partit** à fond **vers la scène de crime**. (Grangé, POL)

-> angl.: *arrive, reach, make one's way to, head to, return to*

# Scène de crime

- Variations sur le collocatif : fr. *lieu(x), parages, environs*
  - (8) Jeanne se souvint que les médias étaient présents **sur les lieux du crime**. (Grangé, POL)
  - (9) Aucun témoin capital **dans les environs du crime**, aucune trace exploitable, et pas le moindre indice sur la lettre, ni empreinte, ni fibre ou marque significative (Chattam, POL)
    - > **non observé en anglais**

**Identification du motif** (extensions syntagmatiques du pivot et variations paradigmatisques sur les différents éléments)

>> **Vmouv** (*arriver, aller, revenir*) + **Prép** (*sur, dans, autour de, (tout) près de, au seuil de, sur le seuil de, dans les environs de, à proximité de, vers*) + **Dét** + *scène de crime*

PIVOT (*crime*) → COLLOCATION (*scène de crime*) → COLLIGATIONS (*sur la SC*) → MOTIF (variations syntagmatique et paradigmatisque)

→ Du local vers le global (Rastier 2011)

# EP désignant de « menus comportements » stéréotypés (réurrence stylistique)

Correspondances FR-ANGL	EP spécifiques au FR	EP spécifiques en ANLG
<p><i>lancer un regard &lt;-&gt; give a look</i></p> <p><i>sa lèvre tremblait</i>            &lt;-&gt; <i>his/her lip trembled</i> , <i>his/her lower lip trembled</i></p> <p><i>consulter sa montre &lt;-&gt; glance at one's watch</i></p> <p><i>se reculer sur sa chaise &lt;-&gt; lean/sit back in one's chair</i></p> <p><i>tirer une chaise &lt;-&gt; pull up a chair</i></p> <p><i>allumer sa cigarette &lt;-&gt; light/get out a cigarette</i></p> <p><i>caresser sa barbe &lt;-&gt; stroke one's beard</i></p>	<p><i>trembler de tous ses membres</i></p> <p><i>trembler de la tête aux pieds / des pieds à la tête</i>  <b>- angl.</b></p> <p><i>passer la <b>main</b> dans les cheveux vs. run <b>a finger</b> through one's hair</i></p> <p><i>fouiller dans + sac</i>            &lt;-&gt; <i>pick up + bag, drop bag to floor, grab bag</i></p>	<p><i>put one's <b>hands</b> in one's pockets</i></p> <p><i>raise + eyebrow</i></p> <p><i>raise one's <b>hand</b> to</i></p>

# *Lancer un regard*

## (récurrence stylistique)

(résultat inattendu)

- Patron lexico-syntaxique 1 : *Lancer un regard* + ADJ (*circulaire, incertain autour de lui, interrogatif vers qn*) :
- Patron lexico-syntaxique 2 : *Lancer un regard* ADJ à N (expansion **syntagmatique**), présence d'un 3<sup>ème</sup> actant **sous forme de N** (comparé à la SF en fr):

(10) **Il lance à Oscar Avane un long regard** d'épagueul pour le remercier de son pieux mensonge (Dard, POL)

> EP fort spécifique au corpus POL fr comparé au corpus litt. général fr (loglike **182**)

- *give (sb) a (ADJ) look* , *give a look at/towards sb/sth* = vers qqn
- > coïncidence des patrons syntaxiques, différence de fréquence et de degré de spécificité: moins spécifique qu'en FR (loglike **32,8**);

(11) Marino gave me a questioning look. (Cornwell)

(12) Hector, curled in the chair opposite her, **gave a disdainful look of disapproval.**



# Fonction(s) textuelle(s) des récurrences stylistiques

- Installer le lecteur dans un horizon de sens familier
- Ancrer le récit dans une constellation **générique** de significations (cf. Dufays 2010: 230-232)

→souvent: les EP servent de cadrage stéréotypique du dialogue:  
comportements parfaitement interchangeables:

**1) comportement:** X se recula sur sa chaise et alluma une cigarette / X se recula sur sa chaise et caressa sa barbe / ...:

**2) discours direct:** - Et vous savez pourquoi?

**3) comportement:** Y haussa les épaules / inclina la tête sur les côtés et dit avec effort / murmura .../

# EP relatives à des descriptions stéréotypées (réurrences stylistiques)

EP spécifiques en ANGL	EP spécifiques en FR
<i>edge of the bed</i>	V (de parole) + <i>entre ses dents</i> -> cadrage du dialogue
<i>palm of sb's hand</i>	<i>sur le crâne</i> : synonyme préféré de „tête“
V (de perception) + <i>out of the corner of his/her eye</i> ( <i>du coin de l'oeil</i> )	<i>au bord du gouffre</i>

→ différences interlinguistiques

(13) She sat **on the edge of the bed**, listening to the shower run. (Hiaasen)

(14) **Out of the corner of my eye**, I caught a glimpse of two of the thugs (Martin)

(15) Les poèmes et les journaux intimes, c'est bon pour les pucelles, **cracha-t-elle entre ses dents**. (Brussolo)

(16) Il porte un blouson blanc, capuche rabattue **sur le crâne**. (Sportès)

# *Ses lèvres trop rouges*

## réurrences stylistiques (descriptions)

- **Motif stylistique de l'excès** : très rare en anglais (*her too-red lips*: moins d'1 occurrence pour 1 milliard de mots)
- Patron lexico-syntaxique : N + *trop* + ADJ couleur (expansion **paradigmatique**) :
  - **POL** : N = noms **de parties du visages ou du corps** (*peau, yeux, bouche, lèvres, paupière, cheveux, dents, joues, mains...*) (spécifique au POL: 11/17 occ.)

(17) Fascinant, dit-il, en léchant **ses lèvres trop rouges**. (Aubert, POL)

- **SF** : N= **éléments ou objets naturels** (6/6 occ.) :

(18) Ce paysage trop calme et trop accueillant suintait l'angoisse par **ses pierres trop blanches, ses feuilles trop vertes, son ciel trop bleu** (Jeury, SF).

→ Surremploi du possessif en fr dans cette structure

# Conclusion

- L'approche **quantitative** (statistique) et la méthodologie d'extraction des ALS a permis de faire émerger des :
- Récurrences **thématiques** (plutôt attendues) :
  - *scène de crime* : plus spécifique en français, comparé à l'anglais
- Récurrences **stylistiques** (plutôt inattendues) :
  - correspondances : *se reculer sur sa chaise , caresser sa barbe*
  - différences: *fouiller dans son sac, lèvres trop rouges*
- Récurrences **structurelles** : structures à dét. poss. en fr (**ses** *lèvres trop rouges*) ou moins de variations sur la prép. en anglais (**at** *the scene of the crime*)
- Variations syntagmatiques et paradigmatisques sur les EP (**motifs**) (de la localisation, l'excès etc): **unités structurantes pour la construction du sens**
  - >> affinité de certaines EP pour certains (sous-)genres, noyau commun partagé par les auteurs francophones et anglophones, mais aussi des spécificités culturelles (? à vérifier sur de plus vastes corpus)

# Conclusion et perspectives

- L'approche qualitative (analyse **fonctionnelle**) > analyser de manière fine les EP statistiquement spécifiques sur le plan *sémantique* (sous-classes sémantiques d'EP, lieu, personnages, comportements) , *syntaxique* (colligations) et *discursif* (fonctions textuelles)
- En perspective, étudier le rôle des motifs, en tant qu' « unités discursives observables et dénombrables » dans l'organisation textuelle des romans (dimension réticulaire; structures phrastiques, relations inter-phastriques, dynamiques temporelles, séquences narratives et descriptives) (Longrée & Mellet, 2013)
- Grâce au repérage automatique des « colligations textuelles » (Hoey, 2005), analyser les préférences ou les aversions des EP pour des positions / fonctions syntaxiques spécifiques dans la phrase et au-delà de la phrase (paragraphe), utile aussi pour les études stylistiques (**exemple: positionnement des descripteurs de « menus comportements » en proximité étroite avec les passages de dialogue**)
- Approche globale pour appréhender la grammaire des genres

## Bibliographie

- Bakhtine, M. (1984.) « Les genres du discours » In *Esthétique de la création verbale*. Paris : Seuil, 263-308.
- Biber D. (2009), “A corpus-driven approach to formulaic language in English. Multi-word patterns in speech and writing”. *International Journal of Corpus Linguistics* 14:3, 275-311. J. Benjamins
- Dufays, J.-L. (2010), *Stéréotype et lecture*. Bruxelles: Peter Lang.
- Hoey, M. (2005). *Lexical priming: a new theory of words and language*, London/New York: Routledge.
- Kraif, O. & Diwersy, S. (2012). “Le Lexicoscope: un outil pour l'étude de profils combinatoires et l'extraction
- Legallois, D., (2012). “La colligation : autre nom de la collocation grammaticale ou autre logique de la relation mutuelle entre syntaxe et sémantique?”, *Corpus* n°11, p. 31-54.
- Longrée, D. & Mellet, S. (2013). “Le motif : une unité phraséologique englobante? Étendre le champ de la phraséologie de la langue au discours”, *Langages* n°189 : 68-80.
- Novakova, Sorba, Kraif « Expressions polylexicales spécifiques dans deux sous-genres littéraires: le roman policier et la science fiction », Colloque international GRATO (Grammaire & Texte), Lisbonne 2-4 juillet 2015
- Rastier F. (2011). *La mesure et le grain. Sémantique de corpus*. H. Champion
- Siepmann, D. (2015), „A corpus-based investigation into key words and key patterns in post-war fiction.“ *Functions of Language* 2. (sous presse)
- Sinclair, J. M. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.