

Soutenance de thèse

**Modéliser la littérature de genre : méthodes computationnelles et évolution
des genres dans la fiction de langue française**

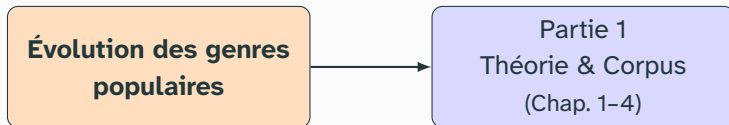
Jean Barré

5 juin 2026

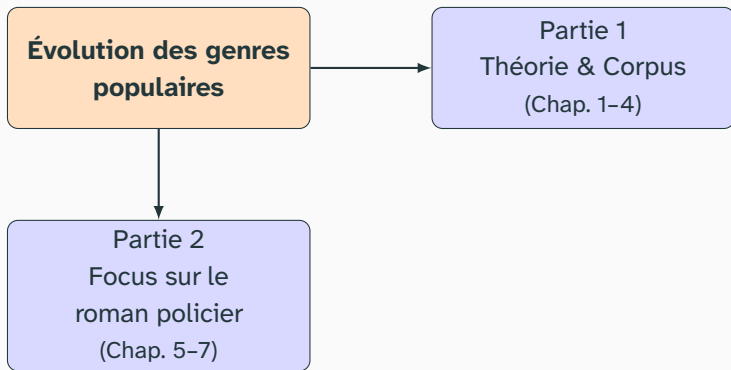
École Normale Supérieure – PSL & LaTTiCe (CNRS, Sorbonne Nouvelle)

Évolution des genres populaires

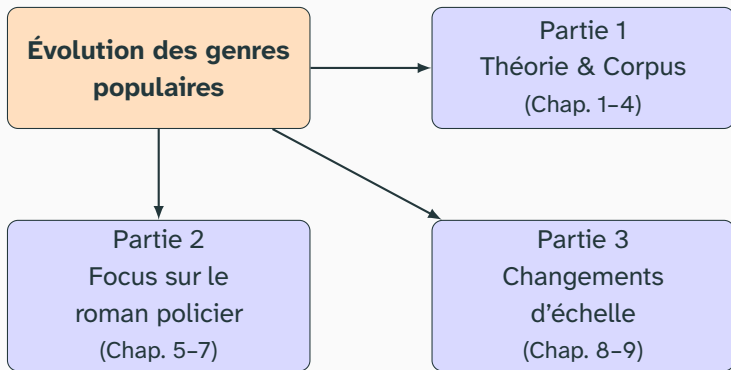
Plan de la thèse



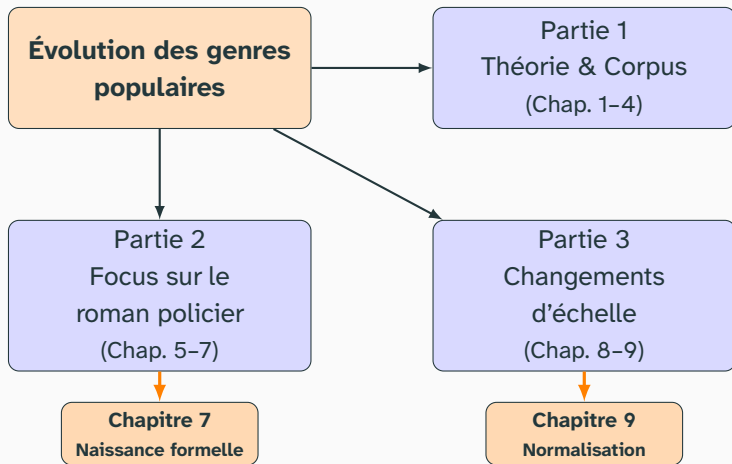
Plan de la thèse



Plan de la thèse



Plan de la thèse



Aujourd'hui : Naissance formelle + Normalisation

Argument central

Les genres populaires ne sont ni de simples étiquettes éditoriales, ni de pures propriétés textuelles.

Argument central

Les genres populaires ne sont ni de simples étiquettes éditoriales, ni de pures propriétés textuelles.

Ce sont des **régularités formelles historiquement situées**, produites par des pratiques sérielles, des attentes de lecture et des dispositifs éditoriaux.

Ces régularités laissent des **traces mesurables** dans les textes – traces que les méthodes computationnelles permettent d'observer à grande échelle.

Les genres populaires ne sont ni de simples étiquettes éditoriales, ni de pures propriétés textuelles.

Ce sont des **régularités formelles historiquement situées**, produites par des pratiques sérielles, des attentes de lecture et des dispositifs éditoriaux.

Ces régularités laissent des **traces mesurables** dans les textes – traces que les méthodes computationnelles permettent d'observer à grande échelle.

1. Naissance (Chap. 7)

Comment un genre se cristallise-t-il
avant sa reconnaissance médiatique ?

Argument central

Les genres populaires ne sont ni de simples étiquettes éditoriales, ni de pures propriétés textuelles.

Ce sont des **régularités formelles historiquement situées**, produites par des pratiques sérielles, des attentes de lecture et des dispositifs éditoriaux.

Ces régularités laissent des **traces mesurables** dans les textes – traces que les méthodes computationnelles permettent d'observer à grande échelle.

1. Naissance (Chap. 7)

Comment un genre se cristallise-t-il
avant sa reconnaissance médiatique ?

2. Normalisation (Chap. 9)

Que devient un genre quand il est reconnu et
éditorialement encadré ?

Argument central

Les genres populaires ne sont ni de simples étiquettes éditoriales, ni de pures propriétés textuelles.

Ce sont des **régularités formelles historiquement situées**, produites par des pratiques sérielles, des attentes de lecture et des dispositifs éditoriaux.

Ces régularités laissent des **traces mesurables** dans les textes – traces que les méthodes computationnelles permettent d'observer à grande échelle.

1. Naissance (Chap. 7)

Comment un genre se cristallise-t-il
avant sa reconnaissance médiatique ?

2. Normalisation (Chap. 9)

Que devient un genre quand il est reconnu et
éditorialement encadré ?

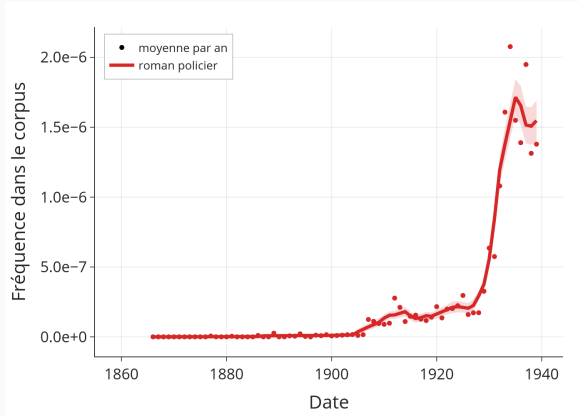
Cas principal : le roman policier • **Comparaison** : le roman d'aventures

Partie 1

La naissance formelle du roman policier

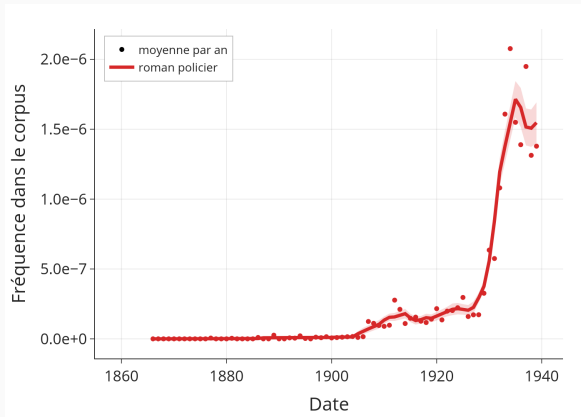
Chapitre 7

La Naissance du Roman Policier



Fréquence de « roman policier » dans la presse française (Gallicagram)

La Naissance du Roman Policier

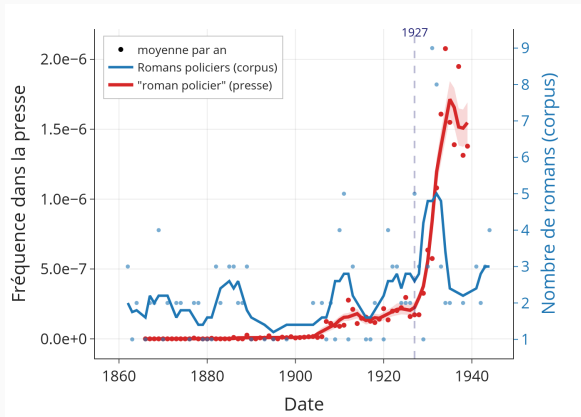


Fréquence de « roman policier » dans la presse française (Gallicagram)

Le paradoxe

- Explosion du terme : **1927**
- Pourtant : ~**plus de 150 romans policiers** publiés avant cette date
- Gaboriau (1832–1873), Leroux (1868–1927), Leblanc (1864–1941), ...

La Naissance du Roman Policier



Fréquence de « roman policier » dans la presse française (Gallicagram)

Le paradoxe

- Explosion du terme : **1927**
- Pourtant : **~plus de 150 romans policiers** publiés avant cette date
- Gaboriau (1832–1873), Leroux (1868–1927), Leblanc (1864–1941), ...

Question : Le genre existe-t-il *formellement* avant sa reconnaissance *contextuelle* ?

Quand est-ce que naît *formellement* le roman policier ?

Quand est-ce que naît *formellement* le roman policier ?

Hypothèse du « génie fondateur »

- « Poe est l'inventeur du roman policier » (Borges, 1936)
- « Gaboriau est le père du roman policier » (Messac, 1929)

⇒ Naissance ponctuelle, auteur unique

Quand est-ce que naît *formellement* le roman policier ?

Hypothèse du « génie fondateur »

- « Poe est l'inventeur du roman policier » (Borges, 1936)
- « Gaboriau est le père du roman policier » (Messac, 1929)

⇒ Naissance ponctuelle, auteur unique

Hypothèse cumulative (Rieder, 2012)

« Étudier les débuts du genre ne consiste pas en une recherche de points d'origine mais en l'observation d'une **accrétion** de répétitions, échos, imitations... »

⇒ Sédimentation progressive

Le roman policier se distingue du roman-feuilleton par trois évolutions structurelles :

1. Persistance du thème criminel

Le crime traverse **tout** le
récit, pas seulement un
épisode.

Le roman policier se distingue du roman-feuilleton par trois évolutions structurelles :

1. Persistance du thème criminel

Le crime traverse **tout** le
récit, pas seulement un
épisode.

2. Centralité du détective

Le détective devient le **pivot**
narratif du récit.

Le roman policier se distingue du roman-feuilleton par trois évolutions structurelles :

1. Persistance du thème criminel

Le crime traverse **tout** le récit, pas seulement un épisode.

2. Centralité du détective

Le détective devient le **pivot narratif** du récit.

3. Structure prototypique

Crime → Raisonnement →
Résolution

Le roman policier se distingue du roman-feuilleton par trois évolutions structurelles :

1. Persistance du thème criminel

Le crime traverse **tout** le récit, pas seulement un épisode.

2. Centralité du détective

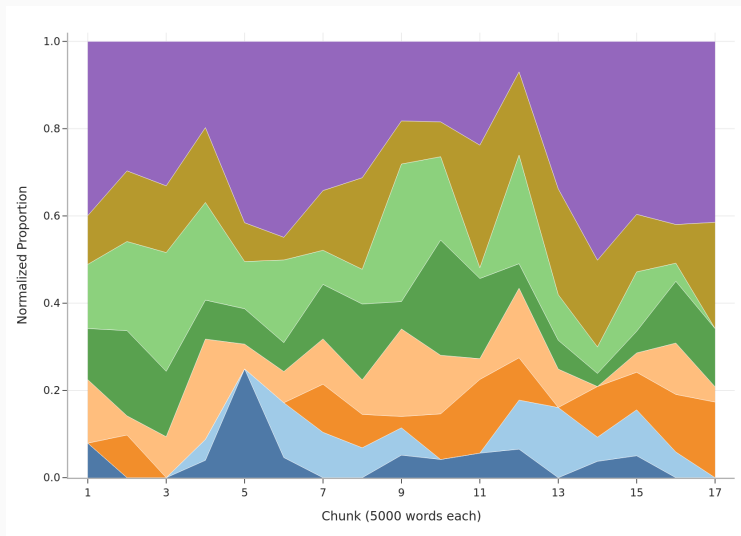
Le détective devient le **pivot narratif** du récit.

3. Structure prototypique

Crime → Raisonnement →
Résolution

Hypothèse : ces trois dimensions cristallisent progressivement et à des moments distincts

Opérationnalisation 1 : Persistance du thème criminel



Mesure de persistance

$$P = 1 - \frac{\sigma}{\mu}$$

- μ : proportion moyenne du thème « Crime »
- σ : écart-type entre morceaux

P élevée $\Rightarrow$ thème stable

Chambre jaune (1907) :

$P = 0.637$

Résultat 1 : Évolution de la persistance thématique

Mesure de persistance

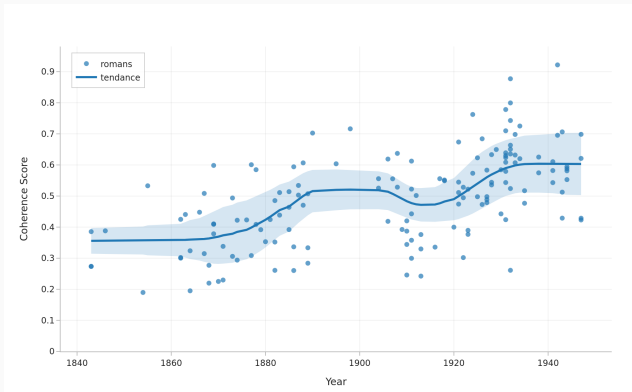
$$P = 1 - \frac{\sigma}{\mu}$$

- μ : proportion moyenne du thème « Crime »
- σ : écart-type entre morceaux

P élevée $\Rightarrow$ thème stable

Chambre jaune (1907) :

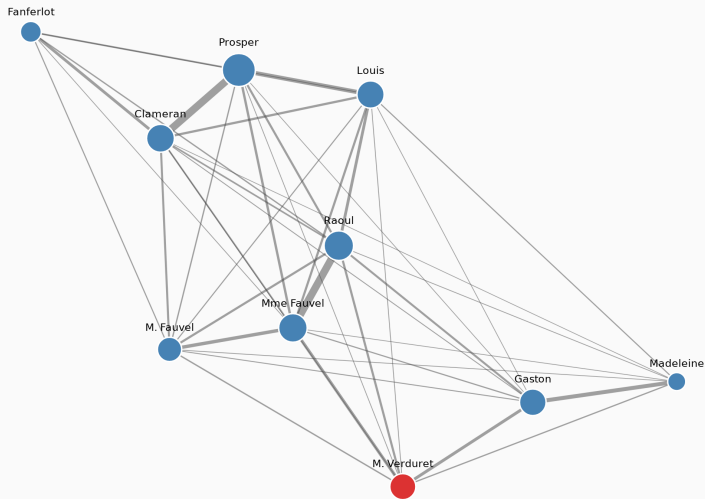
$P = 0.637$



Persistance du thème criminel (1850–1945)

- PROPP-fr :
récupération des
personnages
- Extraction de réseaux
fondés sur les
co-apparitions
- Mesure de centralité :
PageRank

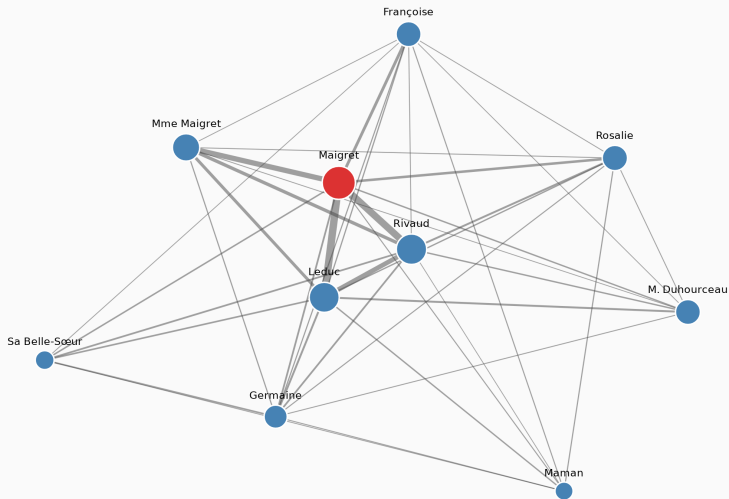
Opérationnalisation 2 : Centralité du détective



Dossier 113
Gaboriau, 1867

PageRank = 0.079

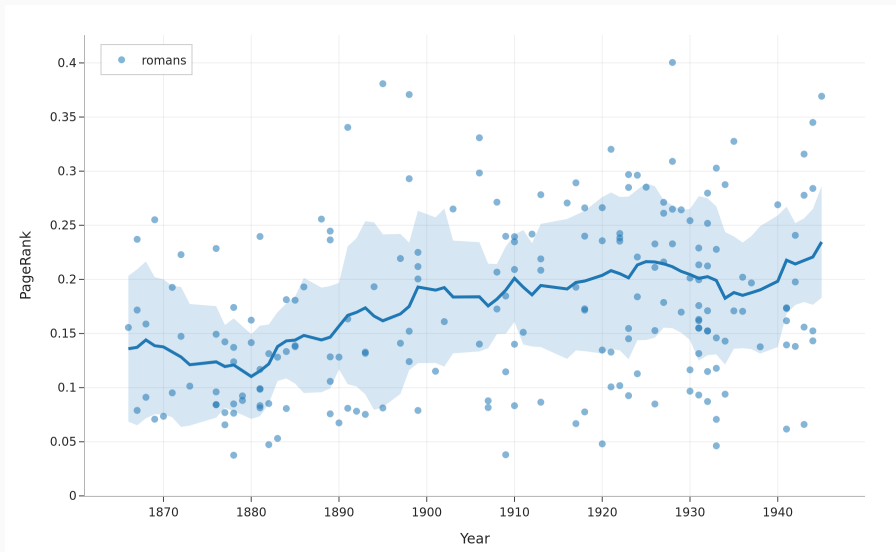
Opérationnalisation 2 : Centralité du détective



Le Fou de Bergerac
Simenon, 1932

PageRank = 0.23

Résultat 2 : Centralité croissante du détective



Évolution du PageRank du détective (1865-1945)

Annotation automatique (Claude Sonnet)

Chaque passage est classé :

- CRIME
- REASONING
- RESOLUTION
- ELSE

$\kappa = 0.83$ (accord inter-annotateurs)

Annotation automatique (Claude Sonnet)

Chaque passage est classé :

- CRIME
- REASONING
- RESOLUTION
- ELSE

$\kappa = 0.83$ (accord inter-annotateurs)

Mesure de prototypicalité (PPlot)

$$\text{PPlot} = 1 - \frac{\text{\#permutations}}{\frac{n(n-1)}{2}}$$

- Compte les inversions par rapport à l'arc canonique
- $\text{PPlot} \in [0, 1]$: 1 = parfaitement prototypique

Annotation automatique (Claude Sonnet)

Chaque passage est classé :

- CRIME
- REASONING
- RESOLUTION
- ELSE

$\kappa = 0.83$ (accord inter-annotateurs)

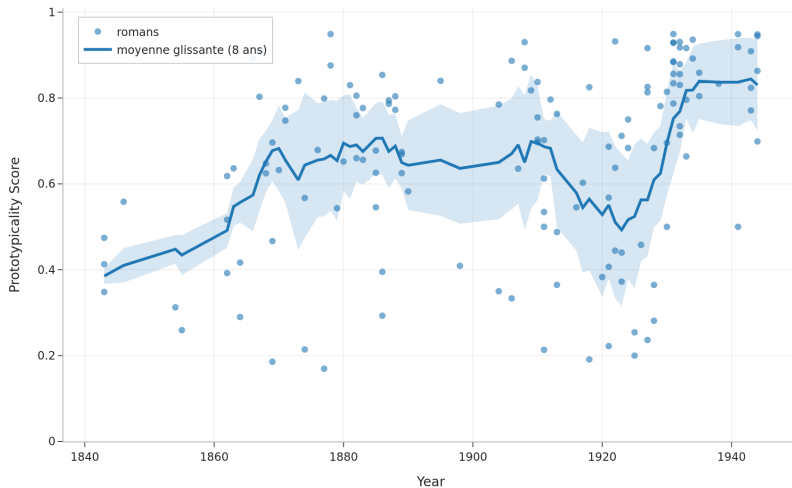
Mesure de prototypicalité (PPlot)

$$\text{PPlot} = 1 - \frac{\text{\#permutations}}{\frac{n(n-1)}{2}}$$

- Compte les inversions par rapport à l'arc canonique
- $\text{PPlot} \in [0, 1]$: 1 = parfaitement prototypique

Arc canonique : Crime $\rightarrow$ Raisonnement $\rightarrow$ Résolution

Résultat 3 : Prototypicalité narrative



Prototypicalité de l'intrigue dans le corpus policier (1860–1945)

Conclusion : trois cristallisations désynchronisées

Persistance thématique

Stabilisation tardive

~ **1930**

Centralité du détective

Convergence progressive

~ **1910**

Structure narrative

Pic de prototypicalité

~ **1935**

Conclusion : trois cristallisations désynchronisées

Persistance thématique

Stabilisation tardive

~ 1930

Centralité du détective

Convergence progressive

~ 1910

Structure narrative

Pic de prototypicalité

~ 1935

Les genres ne naissent pas, ils *cristallisent* progressivement

60 ans de gestation formelle avant la reconnaissance terminologique (1927)

Chapitre 7

Corpus ciblé

Analyse fine des
intrigues, personnages,
thèmes



Chapitre 9

Corpus de projection

Observer les effets de
la normalisation éditoriale
à grande échelle

« *Hors d'usage* » – corpus nettoyé des « *fictiones littéraires* » de Gallica

~ 11 000 romans (1800-1950)

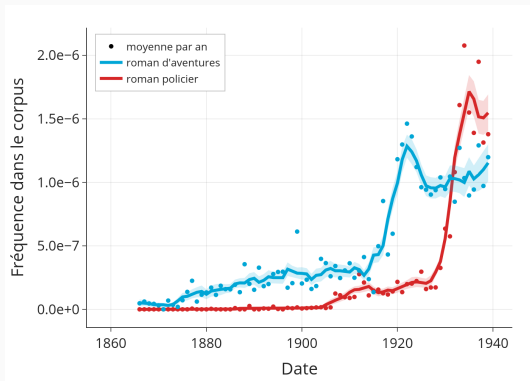
dont 1 375 romans d'aventures & 830 romans policiers

Partie 2

Du feuilleton aux collections : la normalisation des genres

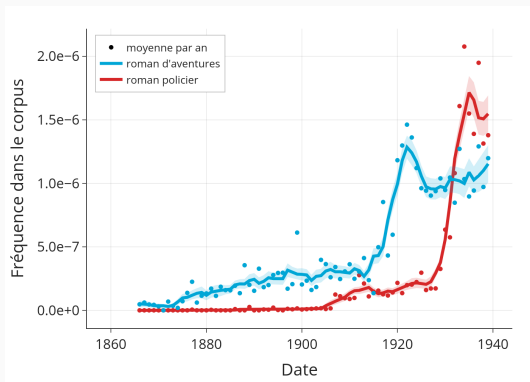
Chapitre 9

Que devient un genre après sa reconnaissance ?



« Roman policier » et « roman d'aventures » dans la presse (Gallicagram)

Que devient un genre après sa reconnaissance ?

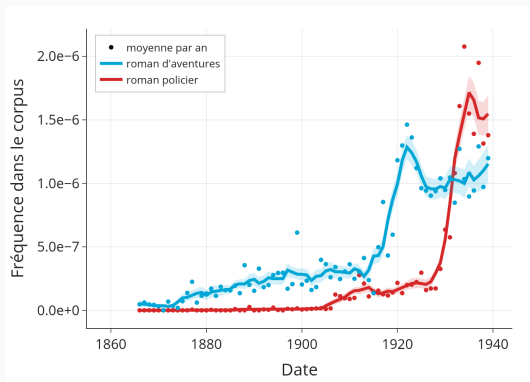


« Roman policier » et « roman d'aventures » dans la presse (Gallicagram)

Letourneux, *Le Roman d'aventures* (2010)

« Cela signifie qu'il existe une rupture dans l'histoire d'un genre, qui tient au moment où il est nommé, isolé. Car si les auteurs s'inscrivent déjà dans une logique stéréotypique avant que le genre ne soit explicitement identifié, **le fait de lui donner un nom tend à accélérer et à systématiser la logique de reformulation sérielle** »

Que devient un genre après sa reconnaissance ?



« Roman policier » et « roman d'aventures » dans la presse (Gallicagram)

Letourneux, *Le Roman d'aventures* (2010)

« Cela signifie qu'il existe une rupture dans l'histoire d'un genre, qui tient au moment où il est nommé, isolé. Car si les auteurs s'inscrivent déjà dans une logique stéréotypique avant que le genre ne soit explicitement identifié, **le fait de lui donner un nom tend à accélérer et à systématiser la logique de reformulation sérielle** »

- **Avant le nom** → Chapitre 7
- **Après le nom** → Chapitre 9

Policier : codification explicite

Van Dine (1928) – *Twenty Rules* pour écrire des romans policiers :

« Règle 3 : *Il ne doit y avoir aucune intrigue amoureuse.* »

Décatalogue de Knox

+ Le Masque (1927), Série Noire (1945)

Policier : codification explicite

Van Dine (1928) – *Twenty Rules* pour écrire des romans policiers :

« Règle 3 : *Il ne doit y avoir aucune intrigue amoureuse.* »

Décatalogue de Knox

+ Le Masque (1927), Série Noire (1945)

Aventures : pas de codification

- « hypergenre » (Letourneux)
- Sous-genres multiples avec des maisons d'éditions spécifiques (Ferenczi, Tallandier)

Policier : codification explicite

Van Dine (1928) – *Twenty Rules* pour écrire des romans policiers :

« Règle 3 : *Il ne doit y avoir aucune intrigue amoureuse.* »

Décatalogue de Knox

+ Le Masque (1927), Série Noire (1945)

Aventures : pas de codification

- « hypergenre » (Letourneux)
- Sous-genres multiples avec des maisons d'éditions spécifiques (Ferenczi, Tallandier)

Hypothèse : deux régimes différents $\Rightarrow$ deux logiques de normalisation ?

Chaque roman est découpé en passages, projetés dans un espace sémantique, puis regroupés en **50 modes discursifs**.

(embeddings + ACP + clustering)

Mode	Mots saillants
Procédural	téléphone, inspecteurs, taxi, commissariat
Judiciaire	procureur, greffier, jury, audience
Néo-polar	flic, boulot, mec, fric
Aristocratie	duc, duchesse, monseigneur, sire
Feuilleton	élisa, pâle, balbutia, larmes

5 des 50 modes du corpus policier

Chaque roman est découpé en passages, projetés dans un espace sémantique, puis regroupés en **50 modes discursifs**.

(embeddings + ACP + clustering)

Une métrique clé

Modes effectifs (K_{eff}) : combien de modes un genre mobilise-t-il réellement ?

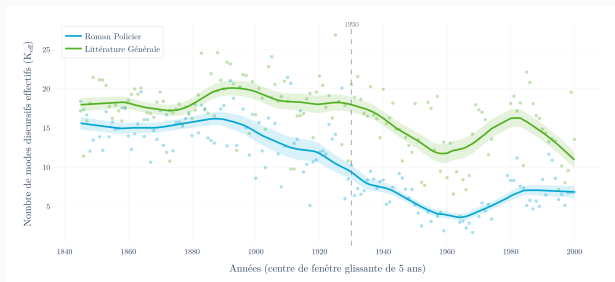
Élevé = grande diversité

Bas = concentration

Mode	Mots saillants
Procédural	téléphone, inspecteurs, taxi, commissariat
Judiciaire	procureur, greffier, jury, audience
Néo-polar	flic, boulot, mec, fric
Aristocratie	duc, duchesse, monseigneur, sire
Feuilleton	élisa, pâle, balbutia, larmes

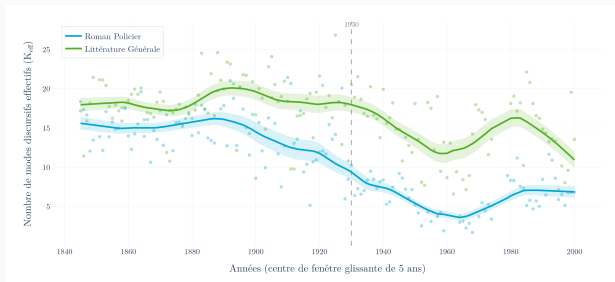
5 des 50 modes du corpus policier

Contraction du roman policier



Modes effectifs (K_{eff}) : policier vs témoin

Contraction du roman policier



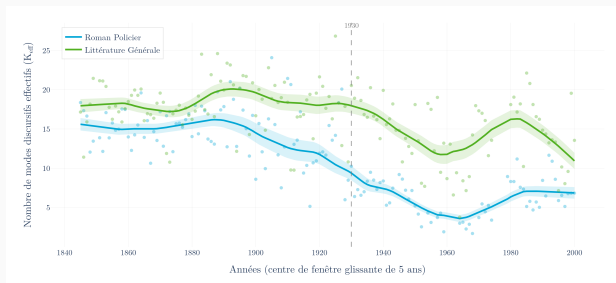
Contraction

K_{eff} : 14.7 $\rightarrow$ 6.3

soit **-57%** après 1930

Modes effectifs (K_{eff}) : policier vs témoin

Contraction du roman policier



Modes effectifs (K_{eff}) : policier vs témoin

Contraction

K_{eff} : 14.7 $\rightarrow$ 6.3

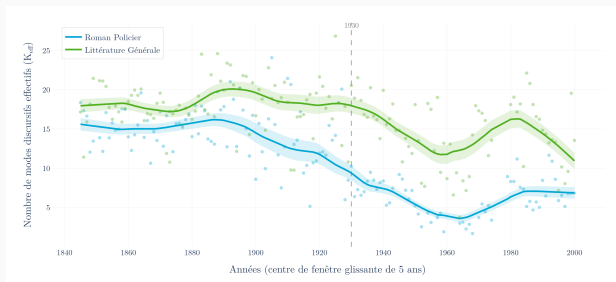
soit **-57%** après 1930

Comparé à un **corpus témoin** (romans non-génériques, même répartition annuelle) :

le policier perd **significativement plus** de diversité.

($DiD = -4.6$, $p < 0.001$; *centroïde stable*)

Contraction du roman policier



Modes effectifs (K_{eff}) : policier vs témoin

Le genre **élimine** ses registres périphériques

Contraction

K_{eff} : 14.7 $\rightarrow$ 6.3

soit **-57%** après 1930

Comparé à un **corpus témoin** (romans non-génériques, même répartition annuelle) :

le policier perd **significativement plus** de diversité.

($DiD = -4.6$, $p < 0.001$; *centroïde stable*)

Lacroix, *Le Banquier de Bristol* (1840)

« Élisabeth était au supplice, les idées les plus noires et les plus désespérantes se heurtaient dans sa tête; son cœur était dévoré d'angoisses; une sueur glacée mouillait ses tempes.

»

Cluster C7

—99% après 1930

- Le corps comme *théâtre* de la souffrance
- L'émotion comparée à la torture physique
- Narrateur omniscient qui exhibe l'intériorité

Disparition du **registre mélodramatique** hérité du roman-feuilleton du xix^e siècle

Simenon, *Maigret et les vieillards* (1960)

« Il attendit l'autobus place Voltaire, eut la chance d'en trouver un à plateforme où il pouvait fumer sa pipe en regardant défiler les rues. Quai des Orfèvres, il salua le planton d'un geste de la main, gravit l'escalier qu'une femme de ménage balayait. [...]

»

Cluster C1

29% de la production post-1930

- Quai des Orfèvres, téléphone
- Gestes du métier
- Le corps comme *indice*, non comme théâtre

Le lecteur reconstruit la psychologie à partir des actes observables.

Du **pathos mélodramatique** à une **écriture behavioriste** (Magny, 1948)

Policier : contraction

Le genre *converge* vers un noyau (l'enquête
procédurale)

Exclusion de modes discursifs obsolètes

Deux logiques de normalisation

Policier : contraction

Le genre *converge* vers un noyau (l'enquête procédurale)

Exclusion de modes discursifs obsolètes

Aventures : fragmentation

Le genre *éclate* en sous-genres (SF, cape-et-épée, exotisme)

Chacun **emporte** une part de l'espace narratif

Deux logiques de normalisation

Policier : contraction

Le genre *converge* vers un noyau (l'enquête procédurale)

Exclusion de modes discursifs obsolètes

Aventures : fragmentation

Le genre *éclate* en sous-genres (SF, cape-et-épée, exotisme)

Chacun **emporte** une part de l'espace narratif

Un point commun

Les deux genres se séparent du **registre mélodramatique du feuilleton**.

Le policier ne peut plus être un roman d'amour accessoirement criminel;
l'aventure, un mélodrame familial accessoirement exotique.

Conclusion

Sur l'histoire des genres

- La dénomination est une **validation** d'un processus formel déjà à l'œuvre, et non une création
- Mais elle **accélère** la spécialisation et la sortie de l'hybridité feuilletonesque
- Selon des *formes opposées* : contraction (policier) vs fragmentation (aventures)

Sur l'histoire des genres

- La dénomination est une **validation** d'un processus formel déjà à l'œuvre, et non une création
- Mais elle **accélère** la spécialisation et la sortie de l'hybridité feuilletonesque
- Selon des *formes opposées* : contraction (policier) vs fragmentation (aventures)

Sur les méthodes computationnelles

- La fiction **longue** met les outils du TAL au défi : adapter, combiner, co-construire
- Ne remplace pas la théorie : la **teste** et la **précise**
- Permet de **dater, désynchroniser, distinguer effet de genre vs époque**

Limites

- Le whodunit comme modèle implicite (roman noir sous-représenté)
- Circularité partielle (lexique → embeddings)
- Corpus francophone, déséquilibré post-1950

Limites

- Le whodunit comme modèle implicite (roman noir sous-représenté)
- Circularité partielle (lexique → embeddings)
- Corpus francophone, déséquilibré post-1950

Perspectives

- Circulation **transnationale** des formes (Moretti)
- Autres genres (anticipation, sentimental)
- Re-diversification post-1970

Merci de votre attention

Discussion

jean.barre@ens.psl.eu • <https://jeanbarre.eu>

La normalisation est-elle un universel ?

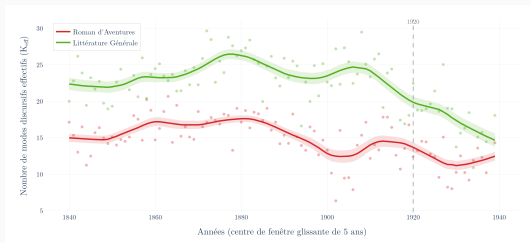
Si la contraction est un effet de la **reconnaissance**, on devrait l'observer dans tout genre reconnu à la même époque.

Test : le roman d'aventures (reconnu dès ~1920, mais *non codifié*)

La normalisation est-elle un universel ?

Si la contraction est un effet de la **reconnaissance**, on devrait l'observer dans tout genre reconnu à la même époque.

Test : le roman d'aventures (reconnu dès ~1920, mais *non codifié*)

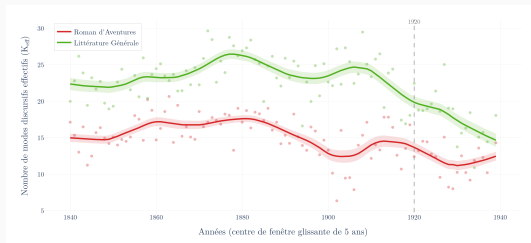


K_{eff} : aventures vs témoin

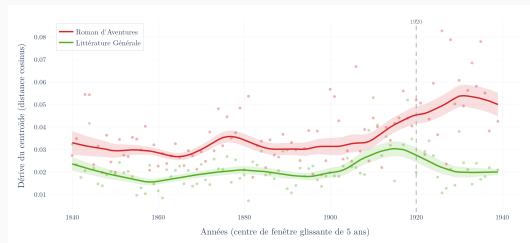
La normalisation est-elle un universel ?

Si la contraction est un effet de la **reconnaissance**, on devrait l'observer dans tout genre reconnu à la même époque.

Test : le roman d'aventures (reconnu dès ~1920, mais *non codifié*)



K_{eff} : aventures vs témoin

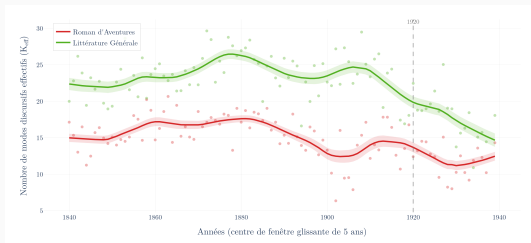


Dérivée du centroïde : aventures vs témoin

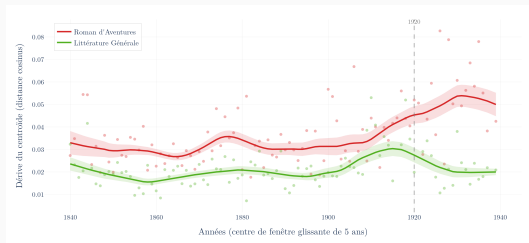
La normalisation est-elle un universel ?

Si la contraction est un effet de la **reconnaissance**, on devrait l'observer dans tout genre reconnu à la même époque.

Test : le roman d'aventures (reconnu dès ~1920, mais *non codifié*)



K_{eff} : aventures vs témoin



Dérivée du centroïde : aventures vs témoin

Pas de contraction spécifique (NS), mais un **déplacement** significatif du centre

Modes qui se contractent

- **Maritime** (−60%)
matelots, frégates, pirates
- **Cape-et-épée de cour** (−57%)
Dumas, Sue : intrigues royales
→ migre vers le roman historique

Le roman d'aventures se fragmente

Modes qui se contractent

- **Maritime** (−60%)
matelots, frégates, pirates
- **Cape-et-épée de cour** (−57%)
Dumas, Sue : intrigues royales
→ migre vers le roman historique

Modes qui s'autonomisent

- **Survie / péril physique** (*50%)
Moselli : évasion, ramping
→ préfigure le thriller
- **Merveilleux scientifique**
Verne, Renard, Rosny
→ devient la SF

Le roman d'aventures se fragmente

Modes qui se contractent

- **Maritime** (−60%)
matelots, frégates, pirates
- **Cape-et-épée de cour** (−57%)
Dumas, Sue : intrigues royales
→ migre vers le roman historique

Modes qui s'autonomisent

- **Survie / péril physique** (*50%)
Moselli : évasion, ramping
→ préfigure le thriller
- **Merveilleux scientifique**
Verne, Renard, Rosny
→ devient la SF

Le genre n'élimine pas, il **redistribue** : chaque sous-genre emporte une part de l'espace narratif.

Le processus de « centrifugation » décrit par Letourneux (2017)

Cluster néo-polar : un second pôle post-1970

Villard, *Rebelles de la nuit* (1987) :

« Sissi et Kader se sont fait sauter la paillasse à la grenade. [...] Sonny Boy qui avait son compte de bidoche malmenée pour la journée ne se rapprocha pas des cadavres. »

flic, merde, boulot, mec, fric, bagnole

Émergence post-1970 (Villard, A.D.G., Vargas)

Héritage du *hardboiled* américain (Cadin, 2018)