Latent Structures of Intertextuality in French Fiction

How literary recognition and subgenres are framing textuality

Jean Barré

December 4, 2024

École normale supérieure - Université PSL

LaTTiCe lab

Structuralist background

Intertextuality

- Kristeva's theory : « Every text is a mosaic of quotations »
- **Explicit** vs. **Implicit Intertextuality**

Textuality

- Literature as an evolving system
- Latent frames of the system : Institutions, Book market, Readers

Aim of the paper - Examine the roles of subgenres and canonization in shaping text similarities

Canon and Genres as Structuring Elements

Canon

- Canonical works as a set of cultural reference points
- Unconsciously defining literary norms - *Cultural grammar* (Altieri, 1983)

Genres

- Genette's *architextuality* i.e. subgenres crystallizing intertextuality
- System of subgenres - an evolving structure that depends on the cultural context

Corpus and Metadata Construction

Corpus Overview

- **Source :** 19,240 texts from Gallica
- Filtered to 12,176 novels with adequate OCR quality
- Spanning 1600 to 1950, with focus on 19th-century texts

Metadata Construction

- Genre labeling with SVM classifier trained on adventure novels
- 1,083 canonical vs. 11,202 non-canonical works

Operationalizing textuality

- SOTA contextual encoder model as a text representation tool
- Fine-tuning M3-Embedding model on French literary language
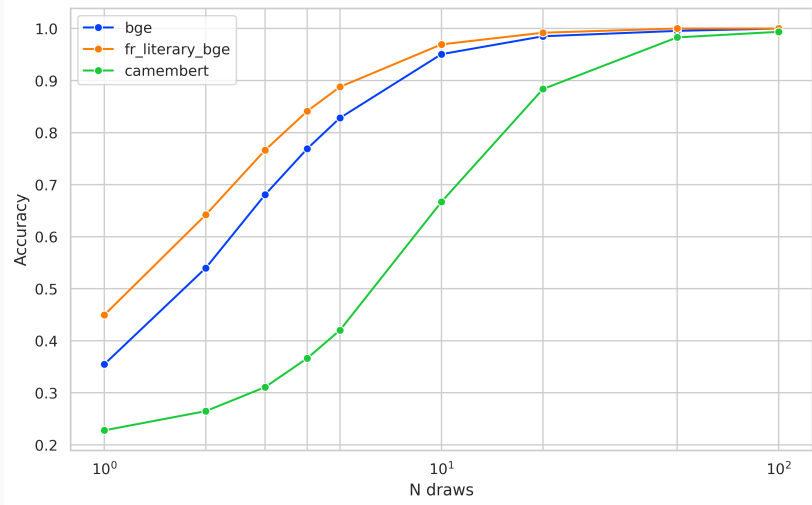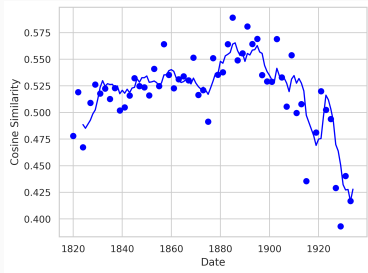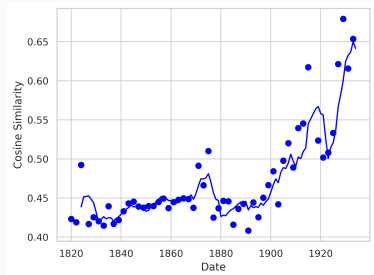
Figure 1 – Sanity check accuracy

Individual trends

(a) Cosine similarity over time for Hugo's *Les Misérables*

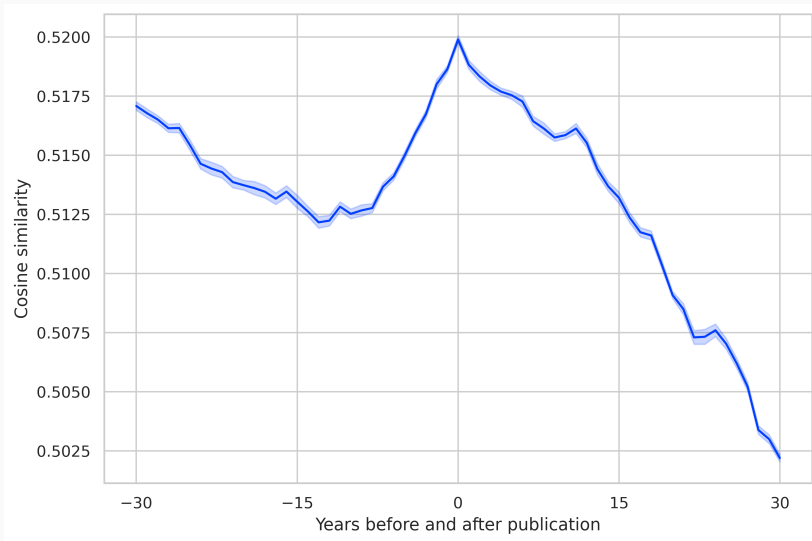**(b)** Cosine similarity over time for Leroux's *Le mystère de la chambre jaune*

Figure 3 – Temporal similarity trends in the corpus

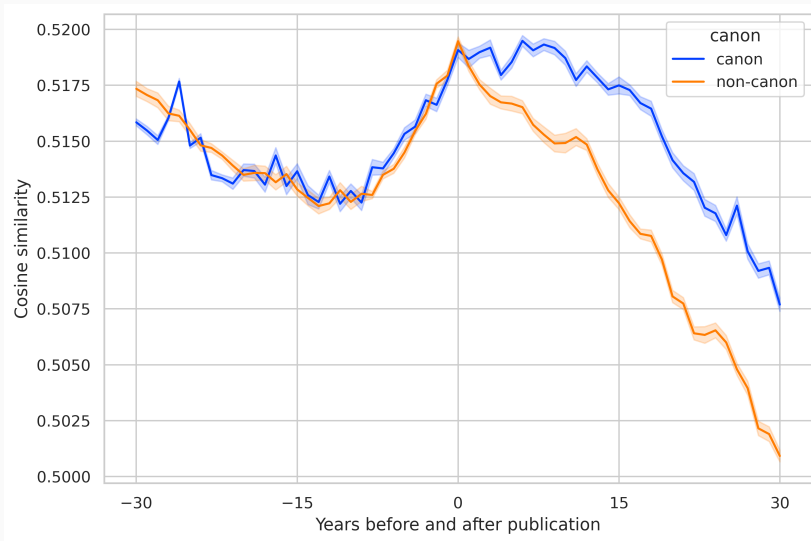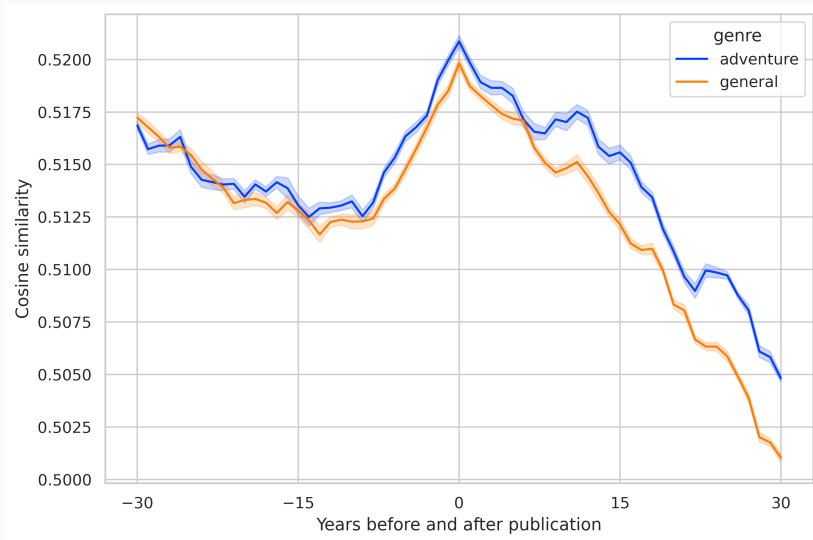**Figure 4 –** Impact of canon on textual similarity over time

Figure 5 – Impact of canon on textual similarity over time

Perspectives

- Canon and Subgenres as underlying frames of textuality
- Intertextuality as a theoretical background for CLS
- Potential for focused sub-genre studies and refined passage-level analysis

Q & A

Feel free to reach me!
jean.barre@ens.psl.eu - jbarre.bsky.social

Data & Code :
```
https://github.com/crazyjeannot/CHR_
         latent_structures
```
Model :
```
https://huggingface.co/crazyjeannot/fr_
         literary_bge_base
```