



The times are a-changin': présent vs passé simple in French novels (1811-2024)

Simon Gabay, Jean Barré, Florian Cafiero

► To cite this version:

Simon Gabay, Jean Barré, Florian Cafiero. The times are a-changin': présent vs passé simple in French novels (1811-2024). DH Benelux 2025, Jun 2025, Amsterdam, Netherlands. hal-04984105

HAL Id: hal-04984105

<https://hal.science/hal-04984105v1>

Submitted on 10 Mar 2025

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.



Distributed under a Creative Commons Attribution 4.0 International License

The times are a-changin’: *présent* vs *passé simple* in French novels (1811-2024)

Simon Gabay
Université de Genève
simon.gabay@unige.ch

Jean Barré
ENS | PSL - LaTTiCe
jean.barre@ens.psl.eu

Florian Cafiero
ENC | PSL
florian.cafiero@chartes.psl.eu

Abstract

The use of *présent* and *passé simple* in French has undergone profound changes in recent centuries. By means of a large corpus of novels, we observe major trends that we attempt to describe and explain. In our results, we discern the birth of the opposition proposed by H. Weinrich between narration (*Erzählung*) and discussion (*Besprechung*), the evolution of which seems to be driven more by literary genres than by the canon.

1 Introduction

According to A. Breton (1924), the poet P. Valéry said that he would never write a novel because he would always refuse to write a sentence like *la marquise sortit à cinq heures* (“the marquise went out at five o’clock”). Among the many poetic problems raised by this sentence (Escola, 2020), we find the use of the *passé simple* for the action verb (*sortit*, eng. “went out”), which reveals a *style d’information pure et simple* (“pure and simple informative style”).

Simultaneously with this change in the perception of the *passé simple*, a revolution is dawning: that of the emergence of the present tense in novels. As G. Philippe (2019) pointed out, this phenomenon is not entirely new, but scholars have observed a clear acceleration at the end of the 19th century, which becomes more pronounced in the 20th century.

Faced with this observation, the question of French verb tenses logically integrates the academic discussion, first with J. Pouillon (1946), and then with the work of É. Benveniste (1959), A. Klum (1961) or H. Weinrich (1964). The latter emphasises the profound divide in French literature between verbal tenses compatible with the *passé simple* (*imparfait* or *plus-que-parfait*), in which the world is narrated (*erzählte Welt*), and those compatible with the *présent* (*passé composé*

or *futur*), in which the world is discussed (*besprochene Welt*). In other words, the question would not be limited to tenses *per se*, but should be studied as a group dynamic, and is not only related to general linguistics, but also to stylistics.

Our paper attempts to revisit this problem of the use of verb tenses in French by describing it quantitatively, and by trying to understand the reasons for the change, using the genre and the canonicity of the works as possible driving forces.

2 State of the art

The question of the evolution of style is a debated topic, whether in traditional stylistics or with more computational approaches, but without anyone from both sides taking into account their respective advances. J. Hughes et al. (2012) computationally demonstrated that textual similarity tends to be temporally localized, with authors showing stronger stylistic similarities to contemporaries than to those from distant periods. More recently, G. Philippe (2021) tried to understand the complexities and reasons for this change over time with a traditional approach.

These modifications of the style over time are strongly linked to language change, which is also extensively studied, particularly in the case of verbs. Recent studies, strongly influenced by evolutionary methods (Lieberman et al., 2007), have emphasised the role of rhymes in this evolution (Newberry et al., 2017). Unfortunately, again, no computational study attempts to truly link such linguistic phenomena with stylistic ones, and *vice versa*.

The question of why change occurs is central. Recent research based on large corpora has identified the importance of genre and canonicity as possible drivers (Barré, 2024), but it may be useful to go back over these two factors in more detail regarding the case of French tenses. In fact, for several years, studies have shown that

changes attributed to canonical authors were more widespread in the literature of their time (Philippe, 2021). It is not impossible that the succession of genres, about which F. Moretti (2005) worked extensively, played a more important role.

3 Corpus

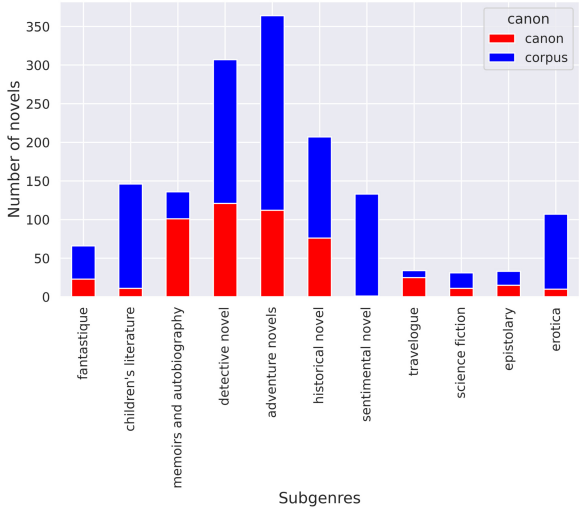


Figure 2: Literary sub-genres in the corpus, broken out by canonicity tags, canon at the author scale

3.1 Data

This study is based on one of the largest corpora for fiction in French: *Chapitres*, a corpus of nearly

3,000 French novels (Leblond, 2022). The investigated period covers over two centuries, from the beginning of the 19th to the 21st century. The corpus is unevenly distributed, both by genre (fig. 2) and chronologically (fig. 3), which introduces potential biases and should be considered when interpreting quantitative results. However, a perfect distribution is not necessarily desirable, the literary production being unstable over time, especially in terms of genre (Moretti, 2005).

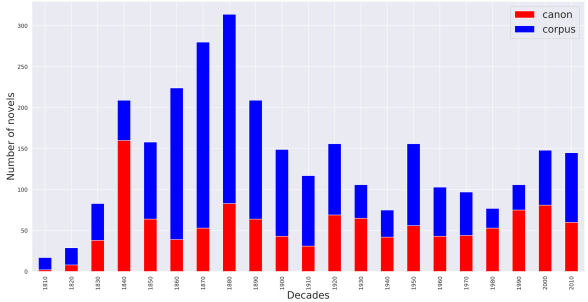


Figure 3: Distribution of the number of novels over time, broken out by canonicity tags, canon at the author scale.

3.2 Metadata

Approximately two-thirds of the *Chapitres* corpus is annotated with subgenre labels. This annotation is based on the classification of the National Library of France (BnF). The validity of these la-

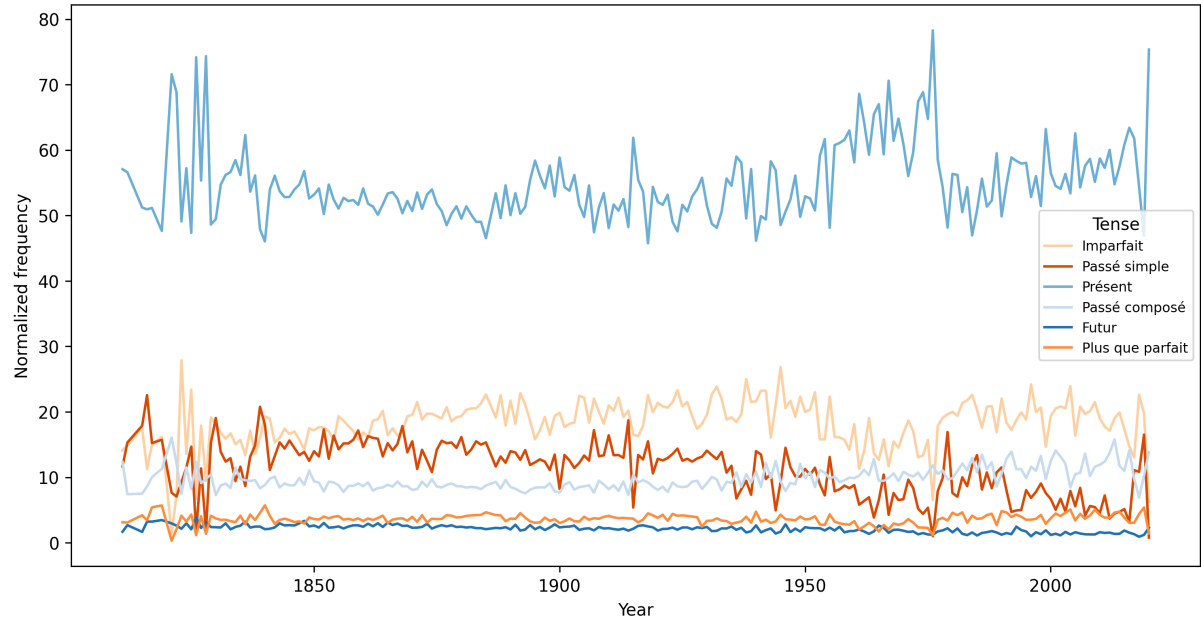


Figure 1: Normalized frequency of verb tenses over the years. Colors distinguish tense groups: warm tones (red, orange) represent past tenses, while cool tones (blue/green) represent present-related tenses. Data at both ends of the graph should be ignored.

bels, discussed by Langlais (2021), remains uncertain, as the practices of the BnF for assigning them have not yet been fully systematised or standardised. Nevertheless, this dataset remains an exceptionally valuable resource for the study of subgenres and is unique in French for its scale and scope.

Regarding the canon, we rely exclusively on previous research (Barré et al., 2023), which focused on canonicity in the context of contemporary reception in French literature. Among the parameters retained, we find in particular: inclusion in school curricula, prestigious collections, literary prizes and other indicators.

4 Experiments

The texts were tokenized and lemmatized using the Stanza toolkit (Qi et al., 2020) with its default French language models.

We identified six French verb tenses - *présent*, *imparfait*, *passé simple*, *passé composé*, *futur*, and *plus-que-parfait* - by analyzing the morphological features provided by Stanza (e.g., Tense and Verb-Form). Compound tenses were identified using grammatical patterns: for example, the *passé composé* was detected through finite auxiliary verbs in the present tense followed by past participles, and the *plus-que-parfait* through auxiliary verbs in the

imparfait followed by past participles.

4.1 First description

A first look at the use of tense usage highlights three notable trends (fig. 1). The first is that, even if they were initially used at comparable frequencies, the *imparfait* emerges as the dominant past tense beginning in the 1870s, at the expense of the *passé simple*. The second is that, shortly before WWII, the *passé composé* becomes more used than the *passé simple*. The third is that some changes are temporary, such as the sudden decrease of the *passé simple* and the *imparfait* in favour of the *présent* in the 1950s.

4.2 Inter-Tense Correlations

Research question A first experiment attempts to determine the relationships between the tenses of French over time, and in particular the past tenses (*passé simple*, *imparfait*, *plus-que-parfait*, in orange) and the present tenses (*présent*, *passé composé*, *futur*, in blue) – following the classification of Weinrich (1964).

Method We compute yearly correlations between the percentage of *passé simple* (reference tense) and the percentage of other tenses. These correlations are calculated per year for all data

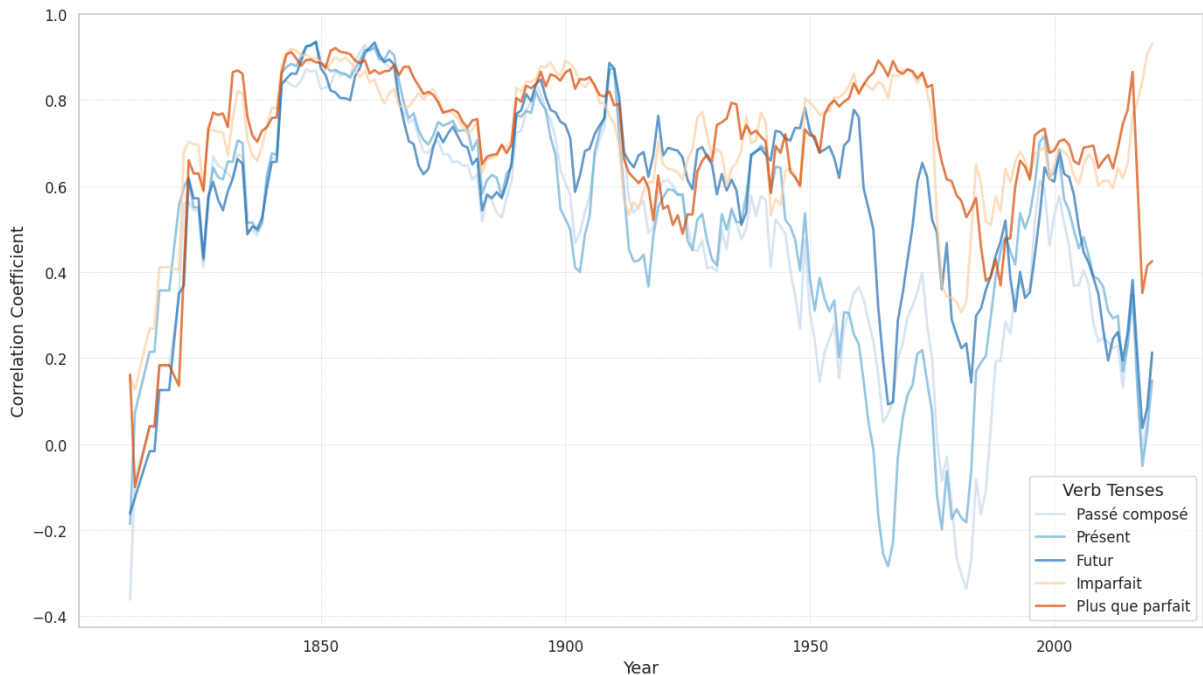
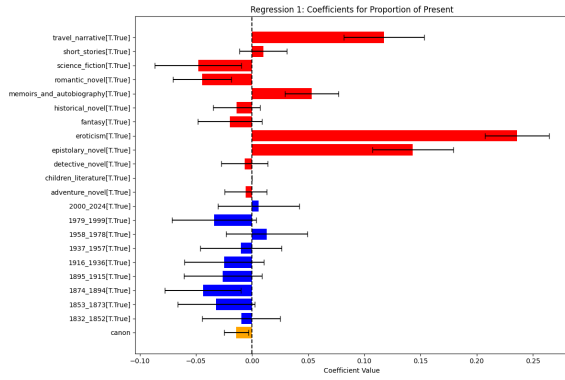
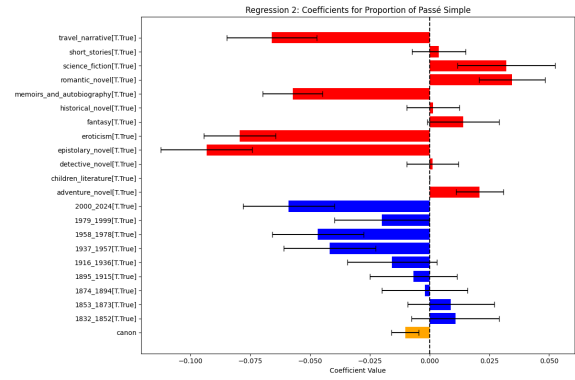


Figure 4: Smoothed Pearson correlations between percentages of use of *passé simple* and of other verb tenses. Once again, colors distinguish tense groups: warm tones (red, orange) represent past tenses, while cool tones (blue/green) represent present-related tenses.



(a) *présent*.



(b) *passé simple*.

Figure 5: Ordinary Least Squares (OLS) regressions.

points available in that year, provided both the reference tense and the compared tense have sufficient variability. A smoothing operation (7-year rolling window) is then applied to mitigate fluctuations and emphasize long-term trends.

Results Fig. 4 displays a smoothed Pearson correlations between the *passé simple* and the other verb tenses. The analysis reveals that after a first phase during which all tenses are strongly correlated, we observe a first limited uncoupling (c. 1900) before a complete decorrelation around WWII with all the other present tenses. However, other past tenses (*imparfait* and *plus-que-parfait*) remain internally correlated with the *passé simple*.

Discussion This first graph confirms not only the increase in usage of the *présent* and corresponding decline of the *passé simple* but also indicates a deeper restructuring process taking place across French verb tenses.. The opposition between present tenses and past tenses is a recent creation, exploding in the 1950s in the wake of the Nouveau Roman, but announced in advance by initial micro-changes. It is also interesting to note that the works of Pouillon (1946) or Weinrich (1964) come just after the decoupling of the present tenses from the past tenses. Their origin is therefore probably linked above all to a reader’s feeling, like Valéry before them.

4.3 Leading edges of tense change

Research question We must now ask ourselves about the causes of this change. G. Philippe (2019) has postulated that “it is not [the *présent*] that has ‘appeared’, it is novelistic protocols that marked a more or less clear break with the realist-naturalist model.” As literary genres influence narrative con-

ventions (e.g., epistolary novels versus historical novels), we aim to quantitatively assess the relationship between genre and changes in tense usage in French literature.

Method We perform two Ordinary Least Squares (OLS) regressions to analyse the proportion of *présent* verbs and to model the proportion of *passé simple*. Both models include canon status, period, and subgenres as predictors.

Results Fig. 5a and 5b indicate that literary genres and historical time periods are key drivers of tense usage changes in French fiction¹. A positive coefficient suggests that a given variable (e.g., genre or time period) is associated with a higher proportion of the given tense, while a negative coefficient indicates an association with a lower proportion. For example, travel narratives, erotica, and epistolary novels strongly favour the present tense, while genres such as science fiction and romantic novels, as well as texts from the late 19th century (e.g., 1874-1894), are associated with significantly lower proportions of the present tense.

Discussion Reading these results, it is clear that it is not the canon that dictates change: it only participates in it, or is even slightly behind. The fact that writers change the tenses they prefer is most likely driven by literary genres, and especially their fictional nature (e.g. epistolary vs romantic novels), which involve linguistic changes with a certain stylistic yield.

¹The pseudo R^2 values for the models hover around 0.4 to 0.5, which suggests that the models explain a reasonable portion of the variance in tense usage.

5 Conclusion

This initial research confirms that the structure of time is undergoing a profound change in the 20th century, and that this change is not being driven by canonical authors. It is a global movement carried out by some literary subgenres, which are most certainly the vehicles of more complex changes in the narrative protocols of the French novel.

References

- Jean Barré. 2024. [Latent structures of intertextuality in french fiction](#). In *CEUR Workshop Proceedings CHR2024*, pages 21–36, Aarhus, Denmark.
- Jean Barré, Jean-Baptiste Camps, and Thierry Poibeau. 2023. [Operationalizing canonicity: A quantitative study of french 19th and 20th century literature](#). *Journal of Cultural Analytics*, 8(3).
- Émile Benveniste. 1959. *Problèmes de linguistique générale*, chapter Les relations de temps dans le verbe français. Gallimard, Paris.
- André Breton. 1924. *Manifeste du surréalisme ; Poisson soluble*. Éd. du Sagittaire, Paris.
- Marc Escola. 2020. [“La marquise sortit à cinq heures” : une anthologie](#). *Atelier de théorie littéraire*.
- James M. Hughes, Nicholas J. Foti, David C. Krakauer, and Daniel N. Rockmore. 2012. [Quantitative patterns of stylistic influence in the evolution of literature](#). *Proceedings of the National Academy of Sciences*, 109(20):7682–7686.
- Arne Klum. 1961. *Verbe et adverbe, étude sur le système verbal indicatif et sur de système de certains adverbes de temps à la lumière ces relations verbo-adverbiales dans la prose du français contemporain*. Almqvist och Wiksell, Stockholm.
- Pierre-Carl Langlais. 2021. [Fictions littéraires de Gallica / Literary fictions of Gallica](#).
- Aude Leblond. 2022. [Corpus chapitres](#).
- Erez Lieberman, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A. Nowak. 2007. [Quantifying the evolutionary dynamics of language](#). *Nature*, 446:713–716.
- Franco Moretti. 2005. *Graphs, Maps, Trees: Abstract Models for Literary History*. Verso, London.
- Mitchell G. Newberry, Christopher A. Ahern, Robin Clark, and Joshua B. Plotkin. 2017. [Detecting evolutionary forces in language change](#). *Nature*, 551:223–226.
- Gilles Philippe. 2019. [Sur l’émergence du présent romanesque](#). *Poétique*, 186:313–329.
- Gilles Philippe. 2021. *Pourquoi le style change-t-il?* Les Impressions nouvelles, Bruxelles.
- Jean Pouillon. 1946. *Temps et roman*. Gallimard, Paris.
- Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. [Stanza: A Python natural language processing toolkit for many human languages](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.
- Heinrich Weinrich. 1964. *Tempus : besprochene und erzählte Welt*. W. Kohlhammer Verlag.