

Rise and Fall of Theatrical Genres in Early Modern France: a Centroid-Based Approach

Anonymous

The 17th c. is the century of the great French tragedy. Corneille and Racine, to name a few, still are among the authors whose notoriety goes beyond the borders of the French-speaking world. Who could have predicted in 1691, when *Athalie* was published, that this major genre of French literature would be, to use Christian Biet’s words, “dead” (Biet 2010, p.167) a century later? Indeed, at the end of the 18th c., tragedy gave way to bourgeois drama with Diderot and Beaumarchais, which would experience its hours of glory in the 19th c. with Hugo (*Hernani*, *Ruy Blas*), Vigny (*Chatterton*) or Musset (*Lorenzaccio*). Conversely, comedy knows another fate: it survives the 18th c. and remains, throughout the 19th c., one of the favourite genres of playwrights.

Many academics have tried to understand this disappearance of tragedy (Allain 1965; Biet 2010...): the rejection of ancient themes by the romantic movement, the incompatibility of tragic values with those carried by the Enlightenment... In contrast, the persistence of the comic genre has not been explained by research. We believe that distant approaches can explain this continuity by revealing a profound internal mutation, partly driven by a more marked presence of the author to the detriment of the rules of the genre.

1 Set up

Studying computationally comedy and tragedy is not trivial. Identifying contents from a specific genre and understanding the rhythm and determinants of these genres’ evolution (Moretti 2005) are much-debated tasks inside the field of Computational Literary Studies. Recent papers on “gothic novels” (Underwood 2016) or “detective novels” (Jockers 2013) demonstrate the complexity of such research.

Depending on context, the detection of the genre of a text relies on different features, be they syntactic (Gianitsos et al. 2019), lexical (Lex, Juffinger, and Granitzer 2010; Al-Yahya 2018), topic-related (Schöch 2017), or on character’s interaction (Hettinger et al. 2015). However, the French Early Modern theatre raises specific problems: because of its often identical sources of inspiration, or its strict compositional rules (be they implicit or explicit), it is characterised by a very strong stylistic and thematic homogeneity compared to the production of the following centuries – a fact that inevitably shows up in computational

analyses (Cafiero and Camps 2019, 2021; Cafiero, Camps, et al. 2020; Gabay 2021). To maximise the linguistic information we get from our texts, we rely on character 3-grams (Kestemont 2014; Sapkota et al. 2015). We choose to keep here the 2,000 most frequent ones as the others were not frequent enough to yield significant effects – but our results remain stable with other values¹. We compute the z-score for all the 3-grams and perform a vector-length Euclidean normalisation (Evert et al. 2017).

To evaluate the heterogeneity within a given set of data compared to a reference situation, we compute centroids for each author and genre. The centroid is defined by the arithmetic mean position of a list of n points ($C = \frac{1}{n} \sum_{i=1}^n x_i$), i.e. in our case the normalised frequency of all the 3-grams of an author or a genre.

Our data come from the *Théâtre classique* corpus (Camps et al. 2021; Fièvre 2007), which transcriptions have all been linguistically normalised (i.e. aligned with contemporary spelling), which simplifies the comparison of plays written at different times, and therefore in different states of the French language. Only the plays belonging to the tragic and comic genres are kept, excluding hybrid genres such as heroic comedy or tragicomedy because they do not offer a clear signal but a mixed one (Schöch 2017). For the calculation of the centroids of the authors, only the authors who have written more than three plays of the same genre. Our *Terminus post quem* is 1550 and our *terminus ad quem* is 1800 (the number of tragedies decreasing drastically thereafter).

2 Experiments

2.1 Experiment A: plays vs the centroid of their genre

By comparing all the plays with the centroid of their respective genre and placing the measured distances of the plays on a chronological axis, it is possible to trace the evolution of comedy and tragedy over time. On fig. 1, the lower the point on the y-axis, the shorter the distance to the centroid: the linguistic practice of the play is more in line with the average of the genre. Conversely, the higher a point, the greater the distance from the centroid: the observed data diverge from the average for the genre. These phenomena can be translated into literary terms: a short distance displays an increased search by the author to conform to the genre, possibly a confinement in it, while a longer distance shows a taking of freedom with the codes.

If we compare the distance of comedies (cf. fig. 1a) and of tragedies (cf. fig. 1b) from their respective centroids over time, we obtain two very different curves. For tragedies, we observe an uninterrupted drop, testifying to ever more similar practices among authors who conform to a common mold. Conversely, for comedies, we observe first an identical trend, and then a clear inflection at the beginning of the 18th c., which reflects a diversification in writing practices.

¹Our code is distributed with this paper for controlling purposes.

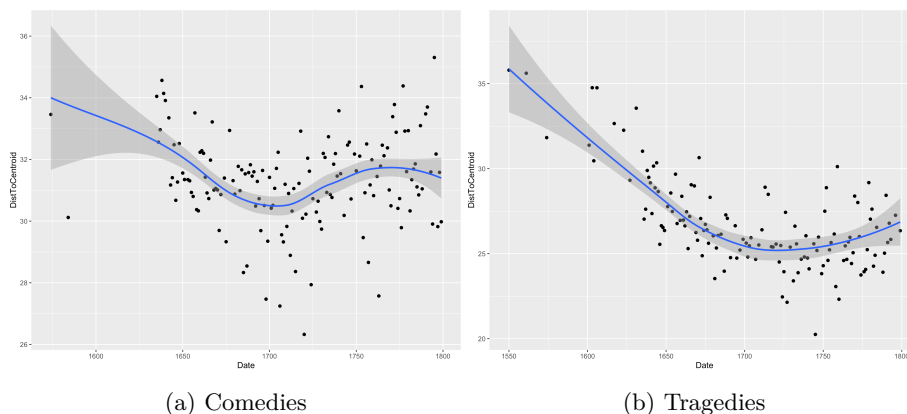


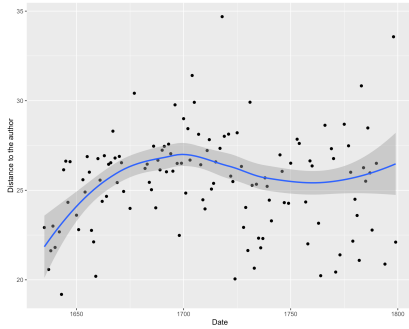
Figure 1: Evolution of the distance from the centroid over time: locally estimated scatterplot smoothing (LOESS). Each point represents a play, the abscissa is that of the date of writing, the ordinate is that of the distance from the centroid.

2.2 Experiment B: auctorial signal vs generic signal

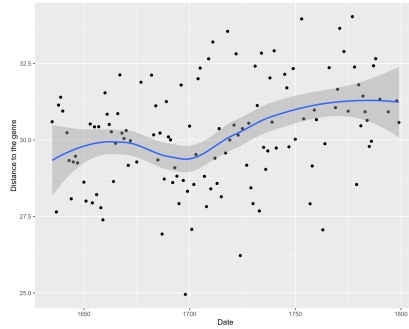
A possible explanation for this change in writing could be found in the clearer affirmation of the personal style of writers, or at least of certain writers, which would temper the attraction exercised by the genre on their work. In order to test this hypothesis, we need to

1. dynamically recalculate the centroid of the genre by using a subset of comedies written 30 years before and after the publication of each play. This prevents comedies that are too old or too recent from exerting too great an influence on the position of the centroid;
2. calculate the distance between each comedy and the centroid of all other comedies written by the same author, by selecting only authors who have written at least two other comedies.

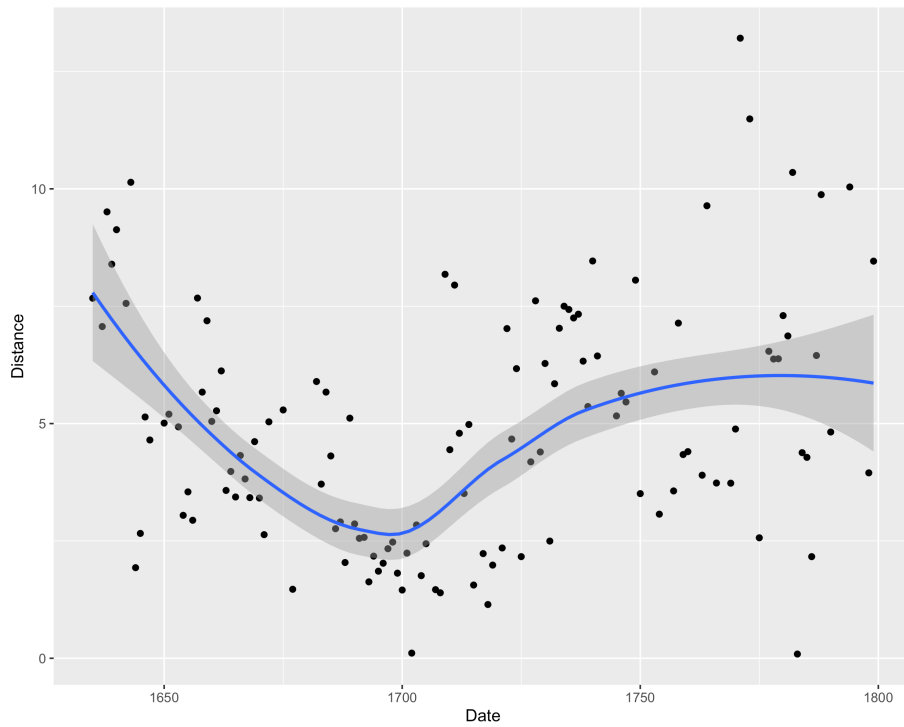
The comparison of the two curves computed with a local regression confirms that in 1700 the distance with the dynamic centroid of the comedies stops falling to increase again (cf. fig. 2b), but it also shows that at the exact same time the distance with the centroid of the comedies of the same author follows an opposite movement: instead of continuing to increase, it begins to stagnate, even to regress slightly (cf. fig. 2a). The calculation of the euclidean distance between the distances with the centroid of comedies and the centroid of comedies written by the same author clearly highlight this divergence in trend (cf. fig. 2c).



(a) Distance between a play and the centroid of other comedies written by the same author.



(b) Distance between a play and the centroid of all comedies written 30 years before and after



(c) Euclidean distance between the two previous distances.

Figure 2: Separation of the authorial and generic signal for the evolution of the comic genre.

3 Discussion

It is possible to distinguish two phases in the history of comedy during early modern France, with a turning point between the 17th and 18th c. Once the codes of the literary genre have been established, comedy manages, unlike tragedy, to halt its fossilisation: by stopping to use common linguistic features, authors reverse the movement of convergence. This liberation of the practice could be done in favour of more personal practices, as evidenced by the curve formed by the distance of each play with the other comedies by the same playwright, which regresses slightly. In other words, insofar as the author is not bound by external generic rules, he can somehow follow his own.

It is important to note that this movement is not clear: we do not find a massive trend, followed uniformly by all the authors. They act in dispersed order: only a non-parametric regression makes it possible to highlight what appears to be a general trend. This growing dispersion of authorial practices is logical: comedy finds itself torn between an old code, still active, and a multiform renewal which is set up in opposition to the old code.

As shown by Jacques Scherer, the genius of 17th century French theater. was, behind seemingly strict rules, to leave writers a margin of freedom (Scherer 2014). This freedom was swallowed up by tragedy, which died of it, and regained by comedy, which continued to exist.

References

- Allain, M.** (1965). “Voltaire et la fin de la tragédie classique française”. In: *The French Review* 39.3, pp. 384–393. URL: <https://www.jstor.org/stable/384752>.
- Biet, C.** (2010). *La tragédie*. Vol. 2e éd. Cursus. Paris: Armand Colin. ISBN: 978-2-200-25909-9. URL: <https://www.cairn.info/la-tragedie--9782200259099.htm>.
- Cafiero, F. and J.-B. Camps** (2019). “Why Molière most likely did write his Plays”. In: *Science advances* 5.11, eaax5489. DOI: 10.1126/sciadv.aax5489.
- (2021). “‘Psyché’ as a Rosetta Stone? Assessing Collaborative Authorship in the French 17th Century Theatre”. In: *Proceedings CEUR-WS* 1613, p. 0073. URL: https://ceur-ws.org/Vol-2989/long_paper51.pdf.
- Cafiero, F., J.-B. Camps, et al.** (2020). “La naissance du style: auteur vs genre aux XVIIe et XIXe siècles”. In: *Humanistica 2020*. Bordeaux, France. URL: <https://hal.science/hal-02577853>.
- Camps, J.-B. et al.** (2021). “Corpus and Models for Lemmatisation and POS-tagging of Classical French Theatre”. In: *Journal of Data Mining and Digital Humanities*. DOI: <https://shs.hal.science/halshs-02591388v2>.
- Evert, S. et al.** (2017). “Understanding and Explaining Delta Measures for Authorship Attribution”. en. In: *Digital Scholarship in the Humanities* 32.suppl 2, pp. ii4–ii16. ISSN: 2055-7671. DOI: 10.1093/llc/fqx023.
- Fièvre, P.** (2007). *Théâtre classique*. URL: <http://www.theatre-classique.fr/>.
- Gabay, S.** (2021). “Beyond Idiolectometry? On Racine’s Stylometric Signature”. In: *Conference on Computational Humanities Research 2021*. Amsterdam, The Netherlands, pp. 359–376. URL: <https://hal.archives-ouvertes.fr/hal-03402994>.
- Gianitsos, E. et al.** (June 2019). “Stylometric Classification of Ancient Greek Literary Texts by Genre”. In: *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. LaTeCH 2019. Minneapolis, USA, pp. 52–60. DOI: 10.18653/v1/W19-2507. URL: <https://aclanthology.org/W19-2507> (visited on 02/06/2023).
- Hettinger, L. et al.** (2015). “Genre classification on German novels”. In: *26th International Workshop on Database and Expert Systems Applications (DEXA)*. Valencia, Spain, pp. 249–253. DOI: 10.1109/DEXA.2015.62.
- Jockers, M. L.** (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press. DOI: 10.5406/illinois/9780252037528.001.0001.
- Kestemont, M.** (2014). “Function Words in Authorship Attribution: From Black Magic to Theory?”. In: *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*. Gothenburg, Sweden, pp. 59–66. DOI: 10.3115/v1/W14-0908.
- Lex, E., A. Juffinger, and M. Granitzer** (2010). “A comparison of stylometric and lexical features for web genre classification and emotion classification

- in blogs". In: *2010 Workshops on Database and Expert Systems Applications*. IEEE. Bilbao, Spain, pp. 10–14. DOI: 10.1109/DEXA.2010.24.
- Moretti, F.** (2005). *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso.
- Sapkota, U. et al.** (2015). "Not all Character N-Grams are Created Equal: A Study in Authorship Attribution". In: *Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: Human language technologies*. Denver, Colorado, pp. 93–102. DOI: 10.3115/v1/N15-1010".
- Scherer, J.** (2014). *La dramaturgie classique*. Hors collection. Paris: Armand Colin. ISBN: 978-2-200-29127-3. URL: <https://www.cairn.info/la-dramaturgie-classique--9782200291273-p-9.htm> (visited on 02/06/2023).
- Schöch, C.** (2017). "Topic modeling genre: An exploration of French classical and enlightenment drama". In: *Digital Humanities Quarterly*. DOI: 10.48550/arXiv.2103.13019.
- Underwood, T.** (2016). "The life cycles of genres". In: *Cultural Analytics 2*. DOI: 10.22148/16.005.
- Al-Yahya, M.** (2018). "Stylometric analysis of classical Arabic texts for genre detection". In: *The Electronic Library* 36.5, pp. 842–855. DOI: 0.1108/EL-11-2017-0236.