Abstract    This contribution attempts to retrace Leo Spitzer’s (1887–1960) famous stylistic reading of the tragedies of French seventeenth-century author Jean Racine (1639–1699) using digital text collections and computational methods of analysis available today. Spitzer’s analysis was first published in 1928 and richly illustrates the manifestations of a “dampening effect” which Spitzer claims is characteristic of Racine’s style and at the same time functions as the signature style of the French Classical period more generally. The contribution uses a mixed-methods approach, combining corpus-based modeling and reading of stylistic patterns with statistical analyses of their distribution. The present attempt to retrace Spitzer’s study not only reveals new insights into Racine’s and the Classical period’s style, but also serves to highlight the respective strengths and limitations of established (non-digital and/or hermeneutical) and computational (digital, algorithmic and/or quantitative) approaches to stylistic analysis and the contrasting notions of style which underpin them.

Keywords    Racine, Spitzer, French literature, style, replication

“Should we strive to obtain the same results or approach the process of replication as a way of knowing?”
(Rockwell 2016)

1. Introduction

This contribution describes an attempt to retrace Leo Spitzer’s (1887–1960) famous stylistic reading of the tragedies of French seventeenth-century author Jean Racine (1639–1699) using digital text collections and computational methods of analysis available today. Spitzer’s analysis, titled “Die klassische Dämpfung bei Racine”
Christof Schöch

(literally, in English: “The Classical Dampening in Racine”), was first published in 1928 and richly illustrates the manifestations of a “dampening effect” which Spitzer claims is characteristic of Racine’s style and at the same time functions as the signature style of the Classical period more generally.

The approach followed here is to take a new look into the “dampening effect” described by Spitzer, that is, the ten abstract phenomena or stylistic effects into which Spitzer divided it and the roughly fifty stylistic patterns or devices he identified in relation to it. 24 of those 50 stylistic patterns could be modeled using search queries applied to richly annotated, digital versions of Racine’s plays. The remaining patterns proved to be too complex and/or too much dependent on context to be modeled at this point. Despite this limitation, this approach permits us not only to reproduce and verify some of Spitzer’s findings, but also to extend his investigation by retrieving additional instances of the relevant stylistic patterns and comparing them to the ones he chose to mention. In addition, the scope of the investigation is enlarged to compare the instances found in Racine’s works with those found in a collection of tragedies contemporary to Racine. Finally, this approach is extended with quantitative methods, in which the quantitative prevalence of the stylistic patterns in Racine’s work and in his contemporaries’ work is compared.

The mixed-methods approach pursued here, combining a close reading of instances with a statistical analysis of their distribution, helps decide whether the dampening effect is characteristic of Racine—and is therefore best described as an authorial style—or, whether it is characteristic, rather, of the French Classical period more generally—and is therefore best described as a period style. Spitzer himself was ambivalent about this, as he calls the key stylistic principle klassische Dämpfung, with reference to a period style, but analyses it by referring exclusively to Racine, essentially conducting a one-author analysis.

The present attempt to retrace Spitzer’s study not only reveals new insights into Racine’s and the French Classical period’s style, but also serves to highlight the respective strengths and limitations of established (non-digital and/or hermeneutical) and computational (algorithmic and/or quantitative) approaches to stylistic analysis and the contrasting notions of style which underpin them. Ultimately, by closely reenacting a previous study, the research presented here also highlights the continuities and differences between established and computational approaches to literature.

2. Context: Repetitive Research

Because it closely reenacts Spitzer’s analysis using digital data and methods, this study falls into the paradigm of repetitive research (as described in Schöch 2023b; see also Schöch et al. 2020). This study repeats earlier work, “in the sense that [it] actively
Repetitive Research

[seeks] to align [its] research questions or hypotheses, [its] datasets and/or [its] methods of analysis, with research practiced and published earlier. This is done with the explicit aim to approximate an earlier study, but conscious also of the fact that perfectly identical repetition is virtually impossible to achieve.” (Schöch 2023b, 374). In the humanities, such repetitive research can often imply that the earlier research one is trying to repeat has been practiced within a non-digital paradigm (i.e., relying on printed sources) and often also in a non-computational paradigm (i.e. relying on digital data, but applying qualitative methods of analysis). Additionally, this study is repeatable, “in the sense that it (typically) makes all the efforts it can to provide the data, code, and explanatory information that make it possible for others, at a later point in time, to perform the same (or very similar) research again.” (Schöch 2023b, 374). In this way, this kind of research is located between past and future: a (never identical) reenactment of past research, and an invitation for (never identical) further reenactments in the future. This is done with the conviction, or at least in the hope, that this cycle of repetitions is not a sterile treading in the same place, but a productive, insightful upwards spiral.

Elsewhere, I have situated this kind of repetitive research in the larger context of the reproducibility crisis that has affected not only fields like medicine and psychology, but is increasingly a concern also in artificial intelligence and the digital humanities, in particular in computational literary studies (for an introduction, see Fidler and Wilcox 2021). Also, I have attempted to structure the field of repetitive research using a simple typology that explains the various forms this kind of research can take (Schöch 2023b). In short, the typology describes the relationship between an earlier study and its repetition in terms of three key variables: the research question, the method of analysis (including the implementation of that method), and the dataset used. For each of these variables, a repetitive study can attempt to operate either in the same, a similar, a different, or an unrelated manner as the previous study. The typology is not meant to establish these distinctions in a purely categorical fashion: rather, as research questions, data or methods are never entirely identical or completely different from the earlier study, the extreme points in the typology are meant to open up a multidimensional gradient of practices. Such a typology can have a number of uses: Conceptually, it helps structure the field, establish conceptual distinctions and provide terminological clarity. Pragmatically, it offers guidance on what data, code and documentation need to be included with a given publication if one or the other mode of replication should be supported in the future.

The present study bridges both a considerable temporal gap—almost a century has passed since Spitzer first published his analysis—and a substantial methodological and technological difference: Spitzer relied on a close reading of a print edition of Racine’s works, whereas the present study employs digital tools to analyze a corpus of Racine’s works (in the first part of the study) as well as a larger corpus of Classical French tragedies (in the second part of the study), combining qualitative and quantitative analyses.
The first and most extensive part of the present study focuses on a computational repetition of Spitzer’s previous, non-digital study. With respect to the typology of repetitive research, this part can be described as follows: as pursuing a virtually identical research question or hypothesis; as relying on very similar, though digital, data; and as using a very different, algorithmic, method of investigation. Where Spitzer uses a holistic approach, fundamentally a stylistic analysis guided by the guiding principle of the dampening effect, the present study’s method is based on defining formal patterns to be identified in annotated text to identify instances of relevant stylistic patterns. In the terms proposed by the typology mentioned above, this part of the present study would best be called a re-analysis (of data).

The second part of the study enlarges the dataset considerably, so that it can no longer be deemed to be similar to Spitzer’s original data. In addition (and as a consequence), the research question also shifts away from Spitzer’s exclusive and programmatic concern with Racine to a comparison of Racine with his contemporaries. Finally, the statistical methods of comparing the prevalence of items in two subcorpora are alien to Spitzer’s study as well, so that we end up with a study that uses a different (though not unrelated) dataset, a very different method and a different research question, leading to research that, in the terms of the typology, would be called follow-up research.

In addition, as it is probably true of any example of repetitive research, this study of course adds the meta-level of comparing the methods and results of the previous study to the current study. I consider this to be at least as important as the findings with regards to Racine and his contemporaries. Any repetitive research helps recognize the presuppositions, strengths and weaknesses of both the original and the current study: the usefulness of their underlying datasets, the appropriateness and inherent biases of their methods and procedures; and the strength and interest of their findings with respect to the research questions at hand. I hope, therefore, that a study such as the present one is not only an occasion to think again about Spitzer, Racine, style and the Classical Age, but also a good moment for reflection about the relationship between established (hermeneutical and/or qualitative) stylistics on the one hand, and computational (algorithmic, quantitative) literary stylistics on the other hand.

3. Background: Spitzer and Racine

The period stretching roughly from 1630 to 1715 and known in French as l’âge classique remains one of the most prestigious periods in French cultural and literary history. Its most lasting literary legacy lies without a doubt in the théâtre
Together with Pierre Corneille and Mollière, Jean Racine is among the most famous and consecrated authors of French drama of this period. Between 1664 and 1691, Racine wrote 12 plays, among which were nine tragedies, two religious plays and one comedy. This makes him noticeably less prolific than some of his contemporaries, like Pierre Corneille and Mollière, who each wrote more than 30 plays. However, Racine's plays have become to be perceived as the epitome of Classical theater, with their relatively strict adherence to the *règles classiques*, i.e. the rule-based poetics of French Classicism, with their sharp focus on the tragic conflict and their measured expression of intense emotion.

To a significant degree, this perception has been created and sustained by Leo Spitzer and the many generations of his readers, despite the fact that Spitzer appears to focus primarily on literary style. It is true that the dampening effect that Spitzer postulates is at the center of Racine's work is first and foremost a principle of literary style. However, it becomes, in the eyes of Leo Spitzer, the manifestation of the aesthetic center of Racine's work. Jaubert aptly describes Spitzer's approach as follows: “dégager la forme-sens d’une écriture, le principe signifiant non seulement niché au creux du détail, mais structurant toute l’œuvre, assurant de la microlecture à la lecture globale la transitivité des niveaux” (1970, 26; “to disentangle the meaningful form of a writing style, the meaningful principle that is not only inscribed within any details, but which also structures the entire work, in this way guaranteeing, from a reading at the micro-level to a global reading, the transitivity of the levels”). Indeed, for Spitzer, a thorough analysis of literary style does more than discover the sum of the stylistic parts: it provides access to the underlying literary coherence and unifying principle of a great author's work. At the same time, this unifying principle guides and directs any stylistic reading of a particular passage.

As noted above, and starting with the title in German—“Die klassische Dämpfung bei Racine”—, the dampening effect that Spitzer sees at work in Racine's work is ambivalent with regard to the question of whether it is a period style, a genre style, or an authorial style. However, in his analysis, Spitzer programmatically puts Racine’s

---

1 Several studies on the *théâtre classique* have become classics in their own right: on dramatic theory (Bray 1926) and practice (Scherer 1950). A more rounded and more recent overview is Génetiot (2005).
2 The bibliography on Jean Racine is obviously vast. One may wish to start with a witty biography (Viala 1990) or a survey of critical work (Rohou 2005). Statistically-minded readers may enjoy consulting the very early study in lexical statistics by Bernet (1983).
3 Note that this and all following translations are provided by the author and are purposefully literal rather than elegant.
4 Spitzer's essay first appeared in German in 1928 (Spitzer 1931). This text is cited as *Dämpfung*, in the remainder of this study. It took four decades before it was translated into French in the volume *Études de style* (Spitzer 1970). Since then, however, the *Études de style* have become a reference for many students and researchers alike.
work front and center to illustrate this stylistic phenomenon. For example, after explaining the first stylistic pattern, he states: “Ich brauche nun nur noch dies Stilmittel als für R[acine] charakteristisch an verschiedenen Stellen der Andr[omaque] und der späteren Dramen zu belegen” (Dämpfung, 138; “I now only need to show, in different places of Andromaque and the later plays, that this stylistic device is characteristic for Racine”). Elsewhere, he describes one of the stylistic patterns as “eine ganz typische Stilfigur bei Racine” (Dämpfung, 232; “a stylistic device truly typical of Racine”). At the same time, Spitzer never explicitly claims that it could not also be found in other authors, whether contemporaries or predecessors.

But what exactly does he mean by the dampening effect? Spitzer explains himself that he believes it is best understood as a musical metaphor, something akin to the dampening pedal on the piano: that is, as a device that allows for a muted sound, but which really only becomes effective in its alternation and contrast with passages that are unmuted. Therefore, Spitzer describes the dampening effect in Racine as follows:


“[...] the often soberly subdued, rationally cool, almost formulaic quality of this style, which then often suddenly and unexpectedly changes only for moments into poetic singing and experienced form, whereupon, however, a certain cap of rational coolness quickly subdues the reader’s shyly venturing lyrical self-expression. Racine, what is properly Racine, is neither mere formula nor mere lyrical singing, but the sequence and the interlocking of both elements”.

Again, Spitzer describes this pattern of intermittent dampening effects as the essence of Racine (“das eigentlich Racinesche”). Indeed, the dampened passages only become apparent when they contrast or alternate with other, more lively, more expressive, less formally-regulated passages. Despite this fact, Spitzer almost exclusively focuses on those general stylistic principles and those stylistic patterns that produce, in his view, a dampening effect in Racine’s plays.

How, then, does Spitzer proceed? Right from the beginning, Spitzer explicitly places the focus on Racine and Racine alone:
Ich möchte nun im Folgenden die Dämpfungen im Stil Racines (nicht bloß in Wortstellung, Rhythmus und Reim) verfolgen, nicht bezogen auf seine Vorgänger, wie Voßler anregt, wodurch Racine zu sehr als Satellit anderer Sterne erschie-ne, als vielmehr als Selbst-Stern, als Sterrenkosmos, den ich, wie gewöhnlich die Gegenstände meiner Stilforschung, in sich ruhend sehe (Dämpfung, 136).

“In the following, I would like to trace the attenuations in Racine’s style (not merely in word order, rhythm, and rhyme), not in relation to his predecessors, as Voßler suggests, which would make Racine appear too much as a satellite of other stars, but rather as a self-star, as a stellar cosmos, which I see, as usually the objects of my stylistic research, resting in itself”.

Spitzer is very systematic in his approach: he takes passages from all of Racine’s plays, except the one comedy, into account in his analysis. The examples he cites are broadly, though not evenly distributed, quantitatively, among the plays. He also breaks down the overall stylistic dampening effect into ten stylistic principles and about 50 specific stylistic patterns or devices, although his grouping is not always very explicit. Each of the principles is named, and each of the patterns is described and discussed briefly before being illustrated with excerpts from Racine’s tragedies, for a total of 484 examples. Effectively, one could say that his argumentative approach starts from an abstract idea (the dampening effect) and then breaks it down into ever smaller and more concrete stylistic levels: broader stylistic phenomena defined primarily by their effect and contribution to the overall dampening effect are broken down into particular stylistic patterns, defined on the level of semantics, syntax, rhetorical devices or metric structure and illustrated using individual examples from the plays. Conversely, Spitzer re-connects each individual example, each stylistic pattern and each broader phenomenon back to the overarching idea of the dampening effect, ensuring the unity not only of his argument, but also of Racine’s work. This systematic approach is by no means quantitative and can only be called formalistic to a limited extent, but it does create an excellent starting point for bridging the gap between stylistic close reading, as practiced by Spitzer, and algorithmic reading, as proposed in the present study.

As mentioned before, in his extensive investigation of the stylistic devices employed by Racine to produce the dampening effect, Spitzer groups them into ten stylistic phenomena, or groups of stylistic patterns and devices, based on a shared principle or a comparable effect. The following is a list established for the purposes of the present study:

5 Andromaque (158 examples) and Phèdre (140) are most widely cited by Spitzer, followed by Bajazet (83) and Athalie (50).

6 The alphabetic identifiers for each group have been added for convenience; page references in parentheses are to Spitzer’s essay in the German edition from 1931 and mark the beginning of
— A “Die Entindividualisierung” (Dämpfung, 136; “the de-individualization”)
— B “[Die] Dämpfende, das Unmittelbare des Empfindens abschwächend[e Wirkung]” (Dämpfung, 144; “the attenuating effect that weakens the immediacy of the sensation”)
— C “[Das] Kühl-Abgeschwächte” (Dämpfung, 150; “coolly-weakened”)
— D “[Dass] das Ich sich nicht zu sehr aussinge” (Dämpfung, 151)
— E “[Die] Unpersönlichkeit der Rede” (Dämpfung, 157; “the unpersonal character of the speech”)
— F Die “Konturverwischung” (Dämpfung, 163; “the blurring of the contours”)
— G Die “Abkühlung der lyrischen Temperatur” (Dämpfung, 178; “the cooling of the lyrical temperature”)
— H Retardierende Elemente (Dämpfung, 214; “retarding elements”)
— J “Gemalte Aufregung” (Dämpfung, 216; “painted excitement”)
— K “Formeln” (Dämpfung, 225, term: 244; “formulae”)

These categories are introduced by Spitzer not as theoretically-justified or explicitly-defined categories, but in a rather casual manner. On the one hand, they serve to implicitly structure and group the analyses of the individual stylistic devices; on the other, they of course connect each device to its principle or effect and, in this manner, mark its contribution to the overall dampening effect.

4. Data and Tools

Before proceeding to a closer look at some of the stylistic patterns themselves, a description of the dataset used in the present study is in order. The dataset used is not an exact replica of the texts Spitzer used; in fact, Spitzer does not indicate the edition he used. However, the dataset does contain the same 11 plays (nine tragedies and
two religious plays, but not Racine’s only comedy) that Spitzer used in his study and which were first performed between 1664 and 1691: *Alexandre, Andromaque, Athalie, Bajazet, Bérénice, Britannicus, Esther, Iphigénie, Mithridate, Phèdre* and *La Thébaïde*. An additional set of tragedies from Racine’s contemporaries was also included, for the second part of the study. It includes 38 tragedies by Claude Boyer, Jean Campistron, Pierre Corneille, Thomas Corneille, Philippe Quinault and Nicolas Pradon, all first performed between 1660 and 1695 and all written in verse, just as Racine’s plays.

Digital versions of the texts were used that are available from the excellent *Théâtre classique* platform (Fièvre 2007–2022) in an XML-TEI format (P4) with tacitly modernized spelling. Building on this basis, all texts were annotated using Freeling, NLTK and WordNet in Python, to produce a format that has a token-based annotation that covers both morphological information and semantic information. All annotations have been represented in an XML-TEI format compatible with the TXM corpus analysis tool, with each token represented in a “w” (word) element and each token-level annotation represented as an attribute-value pair on the respective “w” element. See the following somewhat verbose code listing, taken from Racine’s *Bérénice*, for an illustration:

```xml
<s n="13">
  <w n="t13.1" form="Quel" lemma="quel" tag="DT0MS0" pos="det" type="xxx" gen="masculine" num="singular" wnsyn="xxx" wnlex="xxx">Quel</w>
  <w n="t13.2" form="fruit" lemma="fruit" tag="NCMS000" pos="noun" type="common" gen="masculine" num="singular" wnsyn="13134947-n" wnlex="noun.plant">fruit</w>
  <w n="t13.3" form="me" lemma="me" tag="PP1CS00" pos="pron" type="xxx" gen="xxx" num="singular" wnsyn="xxx" wnlex="xxx">me</w>
  <w n="t13.4" form="reviendra" lemma="revenir" tag="VMIF3S0" pos="verb" type="main" gen="xxx" num="singular" wnsyn="02004874-v" wnlex="verb.motion">reviendra</w>
  <w n="t13.5" form="d'" lemma="de" tag="SP" pos="prep" type="xxx" gen="xxx" num="xxx" wnsyn="xxx" wnlex="xxx">d'</w>
  <w n="t13.6" form="un" lemma="un" tag="DI0MS0" pos="det" type="xxx" gen="masculine" num="singular" wnsyn="xxx" wnlex="xxx">un</w>
  <w n="t13.7" form="aveu" lemma="aveu" tag="NCMS000" pos="noun" type="common" gen="masculine" num="singular" wnsyn="06732350-n" wnlex="noun.communication">aveu</w>
</s>
```

11 All data (as well as the code used) is available from the companion repository mentioned above.
The analyses were performed using the TXM corpus analysis tool (version 0.8.1 released in June 2020; see Heiden 2010) for the query step, while custom Python scripts were used for the comparative analysis step.

5. Formally Modeling Spitzer’s Stylistic Patterns

The key methodological challenge resides in the formal modeling of the complex stylistic patterns described by Spitzer. He, for the most part, provides a brief description of the stylistic phenomenon in question and then provides further refinement, illustration and interpretation using examples from Racine’s plays. Although Spitzer is quite precise in his descriptions, he does not employ formal definitions of the stylistic patterns in any systematic way, not does he limit himself to surface phenomena. Indeed, the stylistic patterns he describes often depend on semantics and metrical structure, both locally and with respect to their wider context in a given play.

As a consequence of the corpus-based, quantifying approach pursued in this replication study, Spitzer’s descriptions need to be re-implemented or operationalized in a formal, machine-actionable way. The method of choice for this has been to employ the query language Corpus Query Processor (CQP; see Evert and Hardie 2011) as used by TXM, in order to be able to apply the corresponding queries on the annotated data available in the TXM corpus format. This query language can be described as relying on regular expressions operating not only on the word forms, but on all available annotations, as well as capable of taking structural cues into account (although this was not used here). The basic unit of the query is the token, and each token to be queried is described in terms of a more or less strict filter on one or several of the levels of annotation. Sequences and patterns of tokens can be defined as well. As Spitzer defines both formal and semantic constraints, the semantic annotation level was of particular importance here. Once a given stylistic pattern is modeled in this way as a CQP query matching the annotations provided in the dataset, all matching stylistic patterns can be retrieved, checked, investigated, and counted.

In order to assess the quality of these approximations of Spitzer’s patterns, two parameters have been assessed: Firstly, I have checked whether all of Spitzer’s examples
are included in the results obtained with the query. In information retrieval terms, this very roughly corresponds to checking the recall of the query. Secondly, I have checked whether all the results identified by the query correspond to Spitzer’s definition. This is very roughly equivalent to checking the precision of the query. As Spitzer does not provide an exhaustive list of relevant examples, but only a certain number of illustrative examples, it is not a sign of a bad query if more relevant examples are found relative to the number of examples given by Spitzer. As a consequence, however, no numerical accuracy score has been calculated, but rather a qualitative assessment has been derived from these two checks.

5.1 Example: Erasing of Contours

A first, and very simple example comes from the group of phenomena Spitzer calls “Kontourverwischung” (Dämpfung, 163 / group F; “erasing of contours”). The stylistic pattern in question is called “konturverwischende Vokabeln” (Dämpfung, 165; pattern F2; “contour-erasing lexical items”) and can be understood as a list of lexical items that are used by Racine in place of a more direct item, as in the following two examples:

(1) Malgré tout son orgueil, ce monarque si fier
À son trône, à son lit daigna l’associer. (Bajazet II, 1)
   ‘In spite of all his pride, this monarch so proud /
    To his throne, to his bed deigned to associate her.’

(2) Les dieux m’en sont témoins, ces dieux qui dans mon flanc
Ont allumé le feu fatal à tout mon sang. (Andromaque II, 5)
   ‘The gods are my witnesses, these gods who in my side /
    Have lit the fatal fire to all my blood.’

Here, lit is the term chosen instead of mariage (‘bed’ instead of ‘marriage’) and flanc instead of ventre (‘flank’/’side’ instead of ‘stomach’), although clearly in a metaphorical way. Other examples Spitzer cites are sein (again for ventre), hymen (for mariage) and courroux (for colère). In addition, lien or nœud can be used in the same way (for either marriage or family and/or love relationships). Spitzer notes that these are not just more noble terms, with respect to the bienséances that disdain excessively corporeal expressions, but also much vaguer terms. This case is interesting as well because it is an instance of an explicit stylistics of deviance from a norm, where a near-synonym is preferred, for stylistic reasons, over the more usual term.

This pattern, while simple in appearance, is not trivial to model, because some of the terms mentioned by Spitzer can not only be used in place of a more direct, usual
| Britannicus, 36 | sang, De les victimes vous-même interrogez le flanc | Toujours la branle a d'heureuses primores, De Rone pour |
| Iphigénie, 34 | sang, De les victimes vous-même interrogez le flanc | . De le silence de les vents demandeze -leur la cause. |
| Iphigénie, 73 | crime ? Pourquoi moi-même enfin me déchirant le flanc | . Payer sa folle amour de le plus pur de mon sang |
| Phèdre, 50 | fils qu'une Amazone a porté dans son flanc | . Cel_Hippolyte... PHÈDRE. Ah dieu ! GENONE. Ce reproche |
| Phèdre, 86 | sont timéons, ces dieux qui dans mon flanc | Ont allumé le feu fatal à tout mon sang, Ces dieux |
| Phèdre, 39 | une main sôre Il lui fait dans le flanc | une large blessure. De rage et de douleur le monstre bondissant |
| Phèdre, 44 | Un dieu, qui d'aiguillons pressait leur flanc | poudreux. À_travers les rochers la peur les précipite. L'essieu |
| Thébaïde, 55 | sang, Recherchez-en la source en ce malheureux flanc | . Je suis de tous les deux la commune ennemie, Puisque |
| Thébaïde, 6 | Leur exemple l'anime à te percer le flanc | ; Et toi seule verses de tes larmes. Tous les_autres versent |
| Phèdre, 157 | toute heure entourée, Je cherochais dans leurs flancs | ma raison épargné. D’un incurable amour remèdes impauissans ! En_vain |
| Phèdre, 44 | petit mérite. Crot-on que dans ses flancs | un monstre m’a porté ? Quelles sauvages mois, quelle haine |
| Thébaïde, 11 | seint nous renfermant tous deux, Dans les flancs | de ma mère une guerre intestine De nos divisions lui manqua l’ |
| Andromaque, 48 | parait efflige, Et se plait d’un hymen | si longtemps néglige. Parmi les déplaisirs où son âme se noie |
| Andromaque, 64 | , de quel oeil Hermione peut voir Son hymen | différé, ses charmes sans pouvoir ? PYLADE_Hermione, Seigneur, |
| Andromaque, 34 | pieds jusqu’à votre colère. Vous-même à cet hymen | venez la disposer. Est ce sur un rival qu’il s’ |
| Andromaque, 48 | détestera, qui toute votre vie Regrettant un hymen | tout prét à s’achever, Voudra... ORESTE_C’est pour cela |
| Andromaque, 2 | si vu Pynhus, Madame, et votre hymen | s’apprête, HERMIONE_On le dit. Et de_plus, on vient |
| Andromaque, 5 | de lui. Comptez depuis quel temps votre hymen | se prépare. Il a paré, Madame, et Pynhus so |
| Andromaque, 15 | vous conduis à le temple, où son hymen | s’apprête. Je vous cains de le bandeau préparé pour sa |
| Andromaque, 33 | retrouve en toi. Si d’un heureux hymen | la mémoire l’est chère, Montre à le fil à quel |
| Andromaque, 48 | parie de moi. Fais - lui valoir l’ hymen | , où je me suis rangée ; Dis-lui, qu’avant |
| Andromaque, 87 | vais seule à le temple, où leur hymen | s’apprête. Où vous n’osez aller mériter ma conquête. |
| Andromaque, 55 | la gloire de vous plaire, Achevez votre hymen | , j’y consens. Mais du moins Ne forcez pas mes yeux |
| Andromaque, 7 | ai vu le temple, ou mon hymen | s’apprête, Mener en conquérant sa nouvelle conquête, Et d’ |
| Andromaque, 48 | de douleur le temple retarditée. De leur hymen | fatal troublons l’événement, Et qu’ils ne soient uns, |
| Andromaque, 9 | Grâces bravés en leur ambassadeur Dussent de son hymen | relever la splendeur. Enfin avec transport prenant son diastème, Sur |
| Bajazet, 29 | une superbe loi De ne point à l’ hymen | assujetir leur foi. Parmi tant de beaux qui briguant leur tendresse, |
| Bajazet, 33 | même Amurat ne me proromp jamais Que l’ hymen | dut un jour couronner ses bienfaits. Et moi qui m’espérais |
| Bajazet, 37 | à sa perte assez autorisés Par le fatal hymen | que vous me proposez. Que vous dirai - je enfin ? Maitre |
| Bajazet, 27 | en faire une image si noire ? L’ hymen | de Soliman ternit - il sa mémoire ? Cependant Soliman n’était point |
| Bajazet, 29 | de ma vie. Qu’un servile hymen | feraient l’ignominie. Soliman n’avait point ce prétexte odieux. |
| Bajazet, 32 | , Et sans subir le joug d’un hymen | nécessaire, Il lui fit de son coeur un présent volontaire. |
| Bajazet, 40 | tout prêts. Qui m’offre ou son hymen | , ou la mort infâtile ; Tandis_qua me présis Atalide sensible |
| Bajazet, 25 | flant enfin a ma reconnaissance, D’un hymen | infâtile a formé l’espérance. Moi-même rougissant de sa crédulité, |
| Bajazet, 28 | était point lié, L’offre de mon hymen | l’eut - il tant effrayé ? N’eût - il pas sans regret |

Fig. 1 Results for query F2 in Racine’s plays, using TXM (Schöch, CC BY).
term; in particular, *lit* and *lien* are also used with the literal meaning by Racine and the interpretation of the words strongly depend on their particular context. The following very simple CQP query can be used to retrieve these instances:

```
[lemma="sein|flanc|lit|lien|noeud|hymen|courroux|lien"%c]
```

This query relies only on the isolated lemmata for a list of words. First of all, we can see that these lexical items are all but rare in Racine's works, with a total of 318 instances (Figure 1).¹² The most frequent term is *courroux* (119 instances), followed by *hymen* (87 instances) and *sein* (57 instances).

All of Spitzer's examples are included in the results, but there is quite a number of cases where it is debatable whether Spitzer would have included the instances among his examples: Indeed, it appears from a close reading of the various instances of *flanc*, for example, that this appears to be used in two ways: in a metaphorical way (as in the above example), but also in a literal, physical sense (as in the example below):

(3)  Le roi frappé d’un coup qui lui perce le flanc,
     Lui cède la victoire, et tombe dans son sang. (*Thébaïde*, V, 3)
     ‘The king struck by a blow which pierces his side, /
     Gives up the victory, and falls in his blood.’

### 5.2 Example: Spatial and Temporal Paraphrases

Another example concerns the group of stylistic patterns Spitzer describes as creating a weakened immediacy of sensibility (*Dämpfung*, 216 / group B). The stylistic pattern in focus here consists of spatial and temporal paraphrases (pattern B2), notably using the paraphrases *en ces lieux*, *en ce lieu* or *sur ces bords* instead of the direct locative adverb *ici* as well as the paraphrase *en ce jour* instead of the direct temporal adverb *aujourd’hui* (*Dämpfung*, 147).

(4)  Vous savez qu’en ces lieux mon devoir m’a conduite. (*Andromaque*, III, 4)
     ‘You know that in these places my duty led me.’

(5)  Depuis que sur ces bords les Dieux ont envoyé
     La fille de Minos et de Pasiphaé. (*Phèdre*, I, 1)

¹² For a closer look, see the CSV file corresponding to the F₂ (= group F, pattern 2) query’s results in the folder “analysis” of the project repository.
Fig. 2 Results for query B2 in the Racine plays, using TXM (Schöch, CC BY).
'Since on these shores the Gods sent /  
The daughter of Minos and Pasiphae.'

(6)  [...] et je puis, dès ce jour,  
Accomplir le dessein qu’a formé mon amour. (Bajazet, II, 1)  
‘and I can, from this day /  
Accomplish the plan which my love formed.’

The corresponding CQP query is only slightly more complex than in the first example, and similarly doesn’t even require any linguistic or semantic annotation, due to the limited range of word forms involved:

```
[word="de|en|sur"%c][word="ce|ces"%c][?][word="lieu.?|bords|jour|jours"%c]
```

Essentially, this query defines a sequence of four tokens, each delimited by square brackets, corresponding to either en ces lieux, sur ces bords or en ces jours, with some flexibility for one or several words to appear in front of the final noun. This latter element is added because Spitzer gives one example where an adjective is placed in that position. This query corresponds to 107 instances in all 11 plays (Figure 2).

The most frequent paraphrase is clearly en ces lieux (52 instances). The quality assessment test shows that recall is perfect, with this pattern: all 16 of Spitzer’s examples are included in the instances found. However, precision is not flawless, with a small number of instances where the literal sequence does not form a paraphrase for here or now, as in the verses: “Et l’aspect de ces lieux où vous la retenez, N’a rien dont mes regards doivent être étonnés” (Britannicus, III, 8; ‘And the look of these places where you hold her, has nothing of which my glances must be astonished’).13

Spitzer claims that the literal ici is rather rare in Racine’s plays. A quick search shows 152 instances of ici in Racine’s plays, clearly outnumbering the synonymous paraphrases, although not all of them are amenable to a replacement by one of the paraphrases mentioned by Spitzer. Whether or not this should be considered rare, however, cannot be ascertained without a deeper qualitative analysis as well as a comparison of the ratios between paraphrase and direct expression in Racine’s and his contemporaries’ works.

13 Again, the full table of results can be found in the “analysis” folder in the project repository (B2.csv).
5.3 Example: Das entgrenzende où

As a final example, I would like to discuss another stylistic pattern from the group of contour-erasing devices (group F) that Spitzer calls “das entgrenzende où” (F4; ‘the de-bordering where’) and that he describes as follows:

Zur Konturverwischung trägt auch bei das entgrenzende ‘où’ (“wo”), das besonders gern bei Abstrakten eintritt, bei denen man sich schwer eine Örtlichkeit, einen umzirkten Raum denken kann, in den das Wo eindringen könnte (Dämpfung, 168).14

“The delimiting “où” (‘where’) also contributes to the blurring of contours, which occurs especially readily in abstract nouns for which it is difficult to imagine a locality, a circumscribed space, into which the ‘where’ could penetrate.”

Among the examples for this pattern cited by Spitzer are the following:

(7) [...] pour avancer cette mort où je cours (Andromaque, II, 2)  
‘[...] this death, where I aspire’

(8) Ô toi qui vois la honte où je suis descendue. (Phèdre, III, 2)  
‘[...] this humiliation, where I have sunk to’

(9) Parmi les déplaisirs où son âme se noie
    Il s’élève en la mienne une secrète joie. (Andromaque, I, 1)  
‘[...] this pain, where his soul is drowning’

The key marker mentioned by Spitzer is the relative pronoun où, combined with a noun expressing an abstract concept. Looking at the examples, we can note that these abstract nouns are often emotion words (such as shame, melancholy, joy, happiness, or worry). We can also note that they are always followed by a verb phrase (where I run, where I have descended to, where my soul drowns, in the examples above). This pattern is referenced as pattern F4 in the documentation for this study. Translating this stylistic pattern into a CQP query, we can formulate the following:

14 Engl.: “The delimiting ‘où’ (‘where’) also contributes to the blurring of contours, which occurs especially readily in abstract nouns for which it is difficult to imagine a locality, a circumscribed space, into which the ‘where’ could penetrate.”
This again defines a sequence of tokens: first, a token defined as either a noun that is an emotion word according to the WordNet annotation or that corresponds to one in a list of emotions words missing from the WordNet list but important in the context of seventeenth-century tragedy; then, an optional token: a comma; then, a token formed by the word form où, whether capitalized (for example at the beginning of a verse line) or not, followed by an optional number of unspecified words; finally, a verb. The list of possible alternative nouns is so long because the coverage of WordNet is very limited, especially with regard to seventeenth-century vocabulary such as hymen, nœud, courroux or hyménée.

While Spitzer quotes ten instances of this pattern, this query, when applied to the eleven plays by Racine, yields 29 instances in nine different plays and depending on 14 different supporting nouns (Figure 3). The quality assessment for this query shows that the recall test is perfect, with all of Spitzer’s examples being part of the retrieved set, and that the precision test is equally perfect, with no instance found using the pattern that cannot be considered to match Spitzer’s description of the pattern.¹⁵

Something that Spitzer does not assess is that the most frequently used supporting noun is trouble (nine instances), followed by hymen (four instances). The most common personal pronoun is je (13 instances), followed by vous (six instances). Spitzer mentions that this pattern can be combined with the “de-bordering plurals” (found in patterns A1 and A2), and the supporting noun is indeed in the plural form in five out of the 28 cases, although it could be debated whether these are really all de-bordering plurals.

Due to limitations of space, it is not possible to show more examples from Spitzer, CQP queries, and their results for all 50 stylistic patterns.¹⁶ When taking more of a bird eye’s view, it can be noted that out of the 50 stylistic patterns that Spitzer distinguished, only 24 can be modeled with good or satisfactory accuracy. Several other patterns are too reduced in their formal expression to be identified with any precision, as in the case of the pattern Spitzer describes as the “Entindividualisierung durch den unbestimmten Artikel” (pattern A1; ‘the dis-individualization by the indefinite article’). One instance of this pattern is the following:

¹⁵ For a closer look, see the CSV file corresponding to the F4 query’s results in the folder “analysis” of the project repository.
¹⁶ For the remaining stylistic patterns, the corresponding queries and results can be found in the project repository.
<table>
<thead>
<tr>
<th>Query</th>
<th>Pivot</th>
<th>Right context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Britannicus, 11</td>
<td>amour, où je voulais</td>
<td>amener sa tendresse. Mais ce lien de le sang qui nous</td>
</tr>
<tr>
<td>Thébaïde, 3</td>
<td>bonheur où je n'osais</td>
<td>penser. L'un et l'autre à la fois je vous puis embrasser. Commencez donc,</td>
</tr>
<tr>
<td>Bérénice, 3</td>
<td>bonheur où peut-être il n'ose</td>
<td>plus penser. ARSACE. Ah quel heureux destin en ces lieux vous renvoie</td>
</tr>
<tr>
<td>Andromaque, 41</td>
<td>déplaisirs où son âme se noie</td>
<td>Il s'élève en la mienne une secrète joie. Je</td>
</tr>
<tr>
<td>Iphigénie, 2</td>
<td>honte, où me dois</td>
<td>je cacher ? Orgueilleuse rivale, on l'aime, et tu</td>
</tr>
<tr>
<td>Phèdre, 3</td>
<td>honte où je suis désolée</td>
<td>Implacable Vénus, suis-je assez condamnée ? Tu ne saurais plus</td>
</tr>
<tr>
<td>Thébaïde, 60</td>
<td>honte où je me vois</td>
<td>contraint, Et c'est injustement que le peuple me crain</td>
</tr>
<tr>
<td>Andromaque, 48</td>
<td>hymen, où je me suis</td>
<td>rangée : Dis-lui, qu'avant ma mort je lui fus</td>
</tr>
<tr>
<td>Iphigénie, 53</td>
<td>hymen où j'étais destinée</td>
<td>! ARCAS. Le roi pour vous tromper feignait cet hymène.</td>
</tr>
<tr>
<td>Mithridate, 25</td>
<td>hymen où j'évoque</td>
<td>Le roi qui m'attendait au sein de ses Etats, Vt emporter</td>
</tr>
<tr>
<td>Mithridate, 132</td>
<td>malheur où vous êtes plongée</td>
<td>Il te fâche en ces lieux d'abandonner la proie,</td>
</tr>
<tr>
<td>Alexandre, 59</td>
<td>malheurs, où vous n'osez</td>
<td>C'est ce trouble fatal que vous fermez les yeux, Qui ne</td>
</tr>
<tr>
<td>Bérénice, 38</td>
<td>malheurs où le ciel m'a plongé</td>
<td>ne penser. Je connais votre coeur. Vous devez vous attendre Que</td>
</tr>
<tr>
<td>Thébaïde, 27</td>
<td>mélancolie Où j'ai vu</td>
<td>Mon fils est mort, Seigneur. ÉTÉOCLE. Il faut</td>
</tr>
<tr>
<td>Andromaque, 7</td>
<td>mélancolie où j'ai vu</td>
<td>si longtemps votre âme ensevelie. Je craignois que le ciel,</td>
</tr>
<tr>
<td>Andromaque, 10</td>
<td>mort où je cours, Qu'à me dire</td>
<td>une fois ce qu'ils m'ont toujours. Voilà depuis un</td>
</tr>
<tr>
<td>Phèdre, 118</td>
<td>pudeur, Où semble</td>
<td>emblé, Et leur osten de le crime aplani le chemin ;</td>
</tr>
<tr>
<td>Athalie, 72</td>
<td>transports où l'amour vous invite</td>
<td>de son sang reluire la splendeur. Et Dieu par sa voix</td>
</tr>
<tr>
<td>Bérénice, 16</td>
<td>trouble, où flottent</td>
<td>ANTIOCHUS. Arsace, je me vois chargé de sa conduite. Je</td>
</tr>
<tr>
<td>Iphigénie, 3</td>
<td>trouble où je la vois</td>
<td>mes esprits. Je n'avais toutefois à craindre que ses cris</td>
</tr>
<tr>
<td>Bérénice, 4</td>
<td>trouble où je la vois</td>
<td>tomber. J'ai vu devant mes yeux Rome entière assemblée.</td>
</tr>
<tr>
<td>Phèdre, 54</td>
<td>trouble où je suis</td>
<td>HIPPOLYTE. Où tendait ce discours qui m'a glacé d'</td>
</tr>
<tr>
<td>Mithridate, 49</td>
<td>trouble où je suis</td>
<td>MITHRIDATE. Princes, quelques raisons que vous me puissiez dire</td>
</tr>
<tr>
<td>Phèdre, 67</td>
<td>trouble où tu me vois</td>
<td>je ne peux rien pour moi. THÉSEÉ. La fortune à</td>
</tr>
<tr>
<td>Phèdre, 21</td>
<td>trouble où vous jetez</td>
<td>, Soleil, je te viens voir pour la dernière fois,</td>
</tr>
<tr>
<td>Mithridate, 52</td>
<td>trouble où vous jetez</td>
<td>l'amour de votre père, Le tourment de me perdre,</td>
</tr>
<tr>
<td>Britannicus, 53</td>
<td>trouble où vous voulez</td>
<td>mon âme, Parlez. Ne suis-je plus dans votre souvenir</td>
</tr>
<tr>
<td>Bérénice, 25</td>
<td>trouble où vous voulez</td>
<td>mon âme. Que vous a dit Titus ? ANTIOCHUS. Au nom de</td>
</tr>
</tbody>
</table>

Fig. 3 Results for query F4 in Racine’s plays, using TXM.
(10) Le croirai-je, seigneur, qu’un reste de tendresse
Vous fasse ici chercher une triste princesse? (Andromaque, II, 2)
‘Will I believe him, Lord, that a remainder of tenderness / makes you seek here a sad princess?’

Whether or not the (naturally very widespread) use of un(e) or des in a given verse corresponds to what Spitzer has in mind here is difficult to ascertain formally, as it depends very much on the context, notably on who are the speaker and addressee of the phrase in question. Other patterns depend strongly on the varying degree of metric and rhetorical complexity of verses, something which is also out of scope for the current annotation schema, but which would of course be an interesting expansion for future work. An example of this is the pattern described by Spitzer as “ganz einfache Verse oder Halbverse, die auf eine hochrhetorische Versreihe folgen” (pattern K2, Dämpfung 228; “very simple verses or half-verses following a series of highly rhetorical verses”). Both metrical structure and rhetorical complexity would need to be identified and annotated with a high level of reliability in order to make such patterns automatically identifiable.17

What also becomes clear, however, is that the procedure so far has already some advantages. For example, instead of relying only on the selection of examples given by Spitzer, the procedure used here allows us to identify many more (if not all) matching instances in Racine’s work and as a consequence, get a more precise sense of the range of variants of a pattern and the distribution of the pattern in the plays. However, the key question raised in the introduction, namely: whether the dampening effect is an authorial style (proper only to Racine) or a period style (characteristic of Classical tragedy as a whole) cannot be answered when just looking at Racine’s plays: this requires making a comparison, both qualitatively and quantitatively, between the instances found in Racine with the instances found in tragedies written by Racine’s contemporaries. This is what the next section attempts to do.

6. Expansion: Racine and His Contemporaries

As we have already seen in section 3 of this paper, Spitzer programmatically describes his approach as describing Racine’s style not in relation to other authors, but as a self-sufficient object, to which he adds:

17 For the description of an automatic tool for metrical annotation of French Classical verse, see Beaudouin and Yvon (1996). For a survey of the state of the art in metaphor detection, see Rai and Chakraverty (2021).
The task of presenting these, so to speak, absolute attenuations in their relative strength compared to the models and to understand them historically is left to other fellow researchers.”

So far, this study has respected Spitzer’s focus on Racine alone. However, I would now like to expand the focus of the study beyond the Racine cosmos not so much to his predecessors, but to his direct contemporaries. The text collection used for this is an expanded dataset including the eleven plays by Racine as well as the 38 plays by his contemporaries, as described above. The approach will be primarily comparative.

Applying the same search queries to this expanded dataset requires some adjustments. Not all patterns can be transposed directly to the larger dataset, with its wider range of authors and plays. For this reason, most queries were enhanced or ‘generalized’ in this step, for example with additional terms being included among the alternatives. However, these expanded queries were then also used for a second round of analyses on Racine alone and have been incorporated in the results presented above. The aim of this procedure is to avoid, as much as possible, any bias that would lead to results favoring higher number of instances in Racine compared to the contemporaries (any remaining bias is most likely in his favor, though). Also, in this constellation, there is no possibility to do the same kind of quality checks as in the first part of this study, as there is no reference any more to establish the recall. However, the same check for false positives (precision) can of course be made manually.

Using the same queries as above on Racine alone for illustration, we can observe some relevant effects. Taking the case of pattern F4 from the analysis above as an example, it can first of all be noted that there are indeed some additional nouns of emotion that appear in relevant patterns, notably maux and désespoir, but also joie and ennui and several others. However, there is also a small number of false positives.

Spitzer argued against this kind of approach, which is incompatible with his contention that Racine’s work should be treated as a unique and unified stylistic object: “Es läßt sich also streng genommen kein Zug der Sprache des Dichters isolieren und mit parallelen Zügen der Sprache anderer Dichter, die ebenfalls aus ihrem Kontext isoliert werden, vergleichen; die einzelnen Züge einer Dichtung sind vorerst miteinander zu vergleichen, als Glieder, Elemente, Träger eines Systems, einer in sich ruhenden Einheit.” (Dämpfung, 257; “Strictly speaking, therefore, no trait of the poet’s language can be isolated and compared with parallel traits of the language of other poets, which are also isolated from their context; the individual traits of a poem must first be compared with one another, as links, elements, carriers of a system, a unity at rest in itself”).
In the following example, the sequence of words fits the pattern, but the underlying construction is different:

(11) M’offre un sujet de joie où j’en voyais d’ennui (Pierre Corneille, 
_Agesilas_, 1666)

‘Offers me a subject of joy / where I used to see boredom.’

A second observation is that there is a large number of additional instances, with 166 instances found among the contemporaries in addition to the 29 instances found in Racine, for a total of 195. Relative to the lengths of the plays, it appears there are 1.7 instances of this pattern per 10,000 tokens in Racine’s plays, but 2.1 instances per 10,000 tokens in the contemporaries’ tragedies.¹⁹

This kind of comparative analysis can be performed for all 24 patterns that were successfully modeled for retrieval in the corpus (see Figure 4 for an overview). It turns out that eleven out of the 24 patterns are over-represented in Racine, but only in one case is that difference statistically significant. The remaining 13 patterns are under-represented in Racine, but only in four cases is that difference statistically significant. This means that 19 out of 24 patterns (or 79 percent) do not vary in prevalence between Racine and his contemporaries with any statistical significance.²⁰

From this second step of the analysis, it can be concluded that the stylistic patterns producing the dampening effect are by no means exclusive to Racine.²¹ Rather, it seems that the patterns Leo Spitzer identified in Racine’s style are actually quite typical of many tragedies written and performed in the second half of the seventeenth century. In this sense, they appear to be a genre style at the very least. Whether or not they are also prevalent in literary genres other than tragedy in verse, during the same period, or whether they are also similarly prevalent in other periods, is something I need to leave to further research at this stage.

However, if the stylistic patterns that Leo Spitzer detected in Racine’s plays are not specific to his style and vocabulary, what is? There is a substantial tradition in the fields of information retrieval, corpus or computational linguistics, and computational literary studies regarding the extraction of words or other features that are distinctive or characteristic of one group of texts when compared to another group of texts. These features are then often called _keywords_ or _key features_. Such distinctive features can be

---

¹⁹ This means the pattern is clearly under-represented in Racine (with Racine using the pattern only at 66 percent of the level of his contemporaries). However, a Wilcoxon rank-sum test shows that this difference is not statistically significant (p-value = 0.30).

²⁰ Further details can be found in the corresponding folder of the companion repository.

²¹ Of course, this can only be a statement about the 24 patterns modeled here and does not preclude that some or even many of the remaining 26 patterns turn out to be strongly distinctive of Racine.
extracted for various kinds of tokens (for example word forms, lemmata or part-of-speech) as well as for unigrams, bigrams or trigrams, etc.\textsuperscript{22}

\textsuperscript{22} A good introduction to several such measures of keyness is provided by Paquot and Bestgen (2009). An important, more technical evaluation study is Lijffijt et al. (2016). Since John Burrows introduced it in 2007, a keyness measure he called Zeta has received quite a lot of attention in computational literary studies; see Burrows (2007), Craig and Kinney (2009), Hoover (2012), Schöch et al. (2018) and Schöch (2018). This last measure, implemented in the \textit{pyzeta} tool, has been used in the following analysis. See: https://github.com/cligs/pyzeta. The code actually used corresponds to release v0.5.0, 2017 of \textit{pyzeta} (DOI: https://doi.org/10.5281/}

\textbf{Fig. 4} Comparative prevalence of stylistic patterns in Racine’s work and the works by his contemporaries. Ordered by ratio of mean relative frequency. Statistically significant differences are highlighted in green. An interactive version of this graph is available in the companion repository (Schöch, CC BY).
This approach basically turns the approach pursued in the last section on its head: instead of specifying a certain number of stylistic patterns and checking whether they are characteristic of Racine or not, compared to his contemporaries, we now ask: given a certain type of feature definition (like part-of-speech bigrams), which ones among all such features are the most distinctive of Racine, and which ones are the most distinctive of his contemporaries? This also means that we do not start out anymore with a certain stylistic hypothesis and interpretive focus (of the kind: assuming Racine's preference for a muted expression of emotions, which relevant stylistic patterns can we identify in his works?). Instead, we now start only with the assumption of the existence of some difference, and of the applicability of a statistical method, but without a clear hypothesis of what it will bring to the fore. This is all the more true as the results up to now have not suggested a very clear contrast between the two groups of texts under study.

When applying the contrastive analysis using the *pyzeta* implementation, the results for unigrams of word forms, certainly the simplest case, are as described below when focusing on just the nouns (Figure 5).23

As can be seen from the visualization, Racine has a preference for several words related to family relationships (mère ‘mother’, aieux ‘ancestor’, frère ‘brother’). Others can be related to a characteristic vision of the tragic plot (autel ‘altar’, nuit ‘night’, larme ‘teardrop’, cri ‘cry’, but also silence ‘silence’). At least one of these distinctive nouns can be related to a stylistic pattern described by Spitzer, namely pas (‘steps’): it is a key element in pattern F7, the periphrases with a verb like porter ses pas or guider ses pas that are somewhat over-represented in Racine (see above).

Similarly, such a contrastive analysis can be done focusing on the adjectives (Figure 6). In the list of the Racinian adjectives, the balance is clearly on the side of the negative (ennemi ‘hostile’, sévère ‘severe’, infortuné ‘unfortunate’, étranger ‘stranger’, homicidal ‘homicide’, perfide ‘treacherous’, farouche ‘fierce’, odieux ‘hateful’, funeste ‘fatal’, impuissant ‘impotent’) rather than the positive adjectives (superbe ‘superb’, jeune ‘young’, éternel ‘eternal’, content ‘content’, sacré ‘sacred’, libre ‘free’, tranquille ‘quiet’, immortel ‘immortal’). Overall, it should be noted that these lists of words are rather hard to interpret. More abstract features, such as part-of-speech trigrams, tend to be even harder to interpret.

---

23 The parameters used are the following: The Zeta variant sd2 has been used, as it has proven to be particularly robust in Schöch et al. (2018). This Zeta variant differs from the standard Zeta calculation in that a log transformation is applied to the document proportions. Text segments of a length of 3,000 words have been used, which is rather short but appears appropriate given that Classical tragedies typically only have 14–15,000 words.
Fig. 5 Distinctive nouns for Racine (right) compared with his contemporaries (left). Implementation: https://github.com/cligs/pyzeta (Schöch, CC BY).
Fig. 6 Distinctive adjectives for Racine (right) compared with his contemporaries (left). Implementation: https://github.com/cligs/pyzeta (Schöch, CC BY).
7. Conclusion and Future Work

The initial idea of a digital replication of Spitzer’s study of Racine’s style has led from formal operationalization of a close stylistic reading using TXM to rather more quantitative approaches, including comparative and contrastive analyses of tragedies by Racine and his contemporaries.

Regarding Racine and the Classical Age, it has become apparent that the stylistic patterns identified by Leo Spitzer, and which in the aggregate produce a dampening effect, again according to Spitzer, are even more frequent in Racine’s work than Spitzer showed. In addition, the analysis was able to show that these patterns are present in varying degrees in Racine’s work and in the work of his contemporaries, but without very strong (or statistically significant) differences, for the most part. The dampening effect, if understood as the aggregate use of individual stylistic patterns, is therefore clearly not an authorial signature style, but the mark of an entire literary genre during a certain period, at the very least.

In terms of the notion of style that underpins the original analysis of Spitzer and the current reenactment, it can be seen that the two notions are rather different. For Spitzer, style is the unifying principle of Racine’s work that is both a general principle and is manifest in a multiplicity of details that each are seen as deriving from, and at the same time reinforcing, the postulated unifying principle. Spitzer’s analysis is, in this sense, inherently interpretative. At the same time, but more implicitly, the notion of style underlying Spitzer’s analysis is also one of style as a deviation from a norm: what is notable, for a stylistic analysis, is what is different from general expectations. By contrast, the definition of style in quantitative stylistics, as evidenced in the last section of this paper, starts with a statistical operationalization of distinctiveness, in our case via the Zeta measure, and is explicitly contrastive, but does not have a strong hypothesis as to how the stylistic features identified as being characteristic of the target group of text can be interpreted. Another difference is that Spitzer’s notion of style supports and even programmatically requires the focus on a single author. An algorithmic approach to style, and particularly any quantification, however, requires explicit target and comparison domains in order to give meaning to the varying degrees of prevalence of stylistic phenomena.

With regard to the formal modeling of the stylistic patterns identified by Leo Spitzer, this has been successful in only 24 out of 50 cases. To a considerable degree, this is due to limitations in the annotations that I was able to create. For example, it would probably be helpful if the semantic annotation had a much larger coverage.

24 For a more general investigation into notions of literary style and how they have shifted between 1950 and the advent of computational stylistics, see Herrmann, van Dalen-Oskam, and Schöch (2015).
either through an expansion of WordNet or through the use of an appropriate word embedding model.\textsuperscript{25} Similarly, a syntactically parsed text would probably allow for more nuanced analyses of stylistic patterns. Finally, it would be of use for several patterns if metrical annotation could also be used. Providing such improved or additional annotations for the challenging language of seventeenth-century verse tragedies was outside the scope of this study. In part, the limitations are also due to the fact that some patterns defined by Spitzer depend on interpretation in the larger context and are probably hard to model even with more annotations. A more controlled approach to the generalization of the patterns, with additional checks for the appropriate coverage, would also be important. Finally, the extraction of key features has used a rather simple approach based on single word forms. Recent, more sophisticated approaches in phraseology, as exemplified for example in the Phraseorom project, would certainly be worth pursuing in order to achieve the goal of an algorithmic identification of complex and significant stylistic patterns.\textsuperscript{26} An obvious upside of the digital approach to the stylistic patterns is the possibility to transfer a given analysis, with just a little adaptation, to a comparison corpus for further analysis.

These strengths and limitations of the digital approach are the mirror image, in a way, of Spitzer’s original study. He would probably not have agreed neither with the rough operationalizations of the stylistic devices, nor with the quantitative comparison of Racine to his contemporaries that I have practiced here. Spitzer is much more nuanced, flexible and mindful of the semantic and pragmatic context of a stylistic device, when identifying and interpreting relevant passages, compared to the many rough and ultimately often imprecise operationalizations used here. However, Spitzer’s analysis would also be extremely time-consuming to transfer to a corpus of different authors. Not just because all of the material needs to be united, but more importantly, because a principle of unity different from Racine’s would most likely need to be postulated and pursued, if one wanted to do justice both to Spitzer’s approach and the contemporary authors. Finally, Spitzer’s study is not easy to replicate: he never even mentions the edition of Racine’s work that he used, he sometimes describes patterns with precision, but in other instances describes rather their purported effect on readers or provides individual examples in order to explain what he means by a given pattern; finally, he refrains from formulating hypotheses that could be tested algorithmically and statistically.

Taking up this last point, I would like to also return to my initial description of this study as repetitive research. In terms of the repeating aspect, I believe that attempting to repeat or reenact Spitzer’s earlier study has been partially successful, in the sense

\textsuperscript{25} For an introduction to word embedding models used in computational literary studies, see Schöch (2023a).
\textsuperscript{26} See e.g. Kraif (2016), Novakova and Siepmann (2020) and Jacquot, Vidotto, and Gonon (2023, this volume).
that (a) roughly 50 percent of the stylistic patterns Spitzer defined could be transposed into the digital and algorithmic paradigm and (b) a close investigation of all textual instances of these patterns, both in Racine and in the works of his contemporaries, could be conducted. In terms of the repeatable aspect, things are more challenging. Certainly, with the complete set of annotated texts available exactly as used for this study, an important part of the project documentation is present. However, documenting the queries and their results based on these texts in an entirely transparent manner is difficult, because copying the query from the documentation into TXM and saving the results to a file needs to be done manually, something which can easily introduce accidental inconsistencies. Similarly, creating the comparison statistics involves another break in the toolchain, as this is not done directly in TXM but in Python, based on data exported manually from TXM. The whole process would be more transparent, and more directly replicable and executable, if it was all done in the same environment, either by integrating the subsequent analyses into TXM using a Groovy-script and automatically exporting all results and intermediary data, or by performing the initial analyses directly using a script-controlled CQP instance. Such an approach has not been used in the present study, but at least the tools used are either well-established in the community (TXM) or available in the long term (*pyzeta*/*pydistinto*).

Ultimately, I hope and believe that the obvious shortcomings of this study do not induce the conclusion that computational literary studies is a futile enterprise. Rather, I hope that in the near future, someone with the required technical expertise in literary and linguistic annotation and bringing to bear new statistical methods of identifying key stylistic devices will further bridge the gap between the nuanced stylistic close readings of a Leo Spitzer with the scope, speed and scalability of algorithmic approaches.

**ORCID®**

Christof Schöch [https://orcid.org/0000-0002-4557-2753](https://orcid.org/0000-0002-4557-2753)

**References**


