

Conclusion

In 2021, a group of American scholars published in *Cultural Analytics* a study attempting to computationally model ‘minor’ European literatures in light of Pascale Casanova’s theories. Their essay, they argued, “address[ed] the current lack of comparative computational literary studies, with most research still overwhelmingly focused on single national literary frames” (Erlin et al., 2021, 85). Three years later, the newly formed Research Committee on Digital Comparative Literature within the International Comparative Literature Association (ICLA) still acknowledged that “while [...] the development of techniques like computer-assisted distant reading and the digitization of large textual archives have opened new perspectives for the study of literature, the terrain of Comparative Literary Studies remains one of the roughest for digital research”.¹

Treading on such a ‘rough terrain’, this book has tried to employ computational methods to investigate the evolution of European drama during the early modern period. The bid to offer an empirical, data-driven – or better, to use Escobar Varela’s (2021, 7 ff.) more precise terminology, ‘data-assisted’ – perspective on the matter situated it not only within the computational literary studies milieu, but also within the framework of cultural analytics. Accordingly, the analysis followed, albeit not systematically, the blueprint thus laid out by Lev Manovich:

1 See https://icla-dcl.quarto.pub/info/files/ICLA_DCL_Description.pdf.

[...] to observe and analyze culture means to be able to map and measure five fundamental characteristics. The first four are diversity, uniqueness, dynamics (temporal changes), and structure. The last term here means clusters, networks, and other types of relations between many objects — that is, structure as it is understood in exploratory data analysis and unsupervised machine learning, as opposed to in 1960s structuralism. [...] In situations in which artifacts were created using a prescriptive aesthetic or template [...] I can also consider a fifth characteristic: variability. (Manovich, 2020, 215)

Diversity, uniqueness, dynamics, structure, and variability: the previous pages tried, through various proxies and methodologies, to touch upon all five components mentioned by Manovich. In doing so, I have striven to bring together two different approaches: one more goal-oriented, consisting in the re-assessment of the argument put forth in 'Modern European Literature', and another based on the paradigm of exploratory data analysis (Tukey, 1977). I shall now conclude by briefly recapitulating my findings and critically reflecting on the methodology adopted and its shortcomings.

Judging from the results of experiments performed in Chapter Four, Moretti's thesis about the evolution of early modern drama – intended as a consistent movement towards progressive diversification of formal features – stands partially confirmed. On the one hand, the two implementations I adopted for measuring vector distances did not completely agree: the pairwise method showed a convergence-divergence movement, while the centroid-based one rather supported a broad increase in diversification. On the other hand, the inspection of the PCA projections supported more forcefully the narrative of a branching of dramatic traditions; even if the sample per snapshot (30 works) is limited, some clustering does become recognisable.

Finally, the most interesting results come from profiling dramatic spaces according to the evolution of specific metrics. By examining their variation across the different subcorpora, I found traces of persisting shifts in plot structures, characters' features, and network arrangements, suggesting that it is possible to discriminate between dramatic

cultures on the basis of form, and that some of these differences have actually been increasing throughout time. The reproduction experiment on the English and French DraCor collections, while not statistically validating the evidence from EmDraCor, still suggested that it retains value in pointing out the general direction in which the different traditions were developing at the time.

Together, the three approaches proposed – based on the observation of distances, clusters, and patterns – yielded some insights into European drama which go beyond the assessment of Moretti's thesis. The PCA, for example, shows a process which is better described as specialisation than as diversification: dramatic traditions are not expanding into new formal territory, but progressively sorting themselves within a broadly stable space of possibilities, each consolidating a distinct formal identity. Furthermore, computational evidence gives substance to previously qualitative observations, such as the persistent proximity of the French and Italian traditions based on shared poetics. Moreover, several metric shifts correlate with documented historical events, such as the introduction of actresses on the English stage from 1660, showing how formal features meaningfully reflect developments in theatrical practice.

The results of this investigation, however, reveal the limits of applying a single rigid framework, such as speciation, to literary history. While Moretti's 1994 introduction of the 'branching' model within literary studies was pioneering, its emphasis on a straightforward and monodirectional development of drama history has been widely criticised. Tom Eyers, for example, has spoken both of a New Positivist posture and an "implicit Hegelianism" latent in Moretti's theories – recognising in his work a strong teleological drive, based on "the tacit, humanistic assumption that history progresses toward a goal, cultural form perspiring on cue to [...] changes that are to be charted from the perspective of their already having happened, and thus, one may say, from the purview of the victors" (Eyers, 2017, 46–47).

From a disciplinary point of view, it should also be emphasised that, in the thirty years following 'Modern European Literature', many more models of cultural evolution, going far beyond basic speciation, have been proposed. Any attempt at reconstructing the evolution of Euro-

pean early modern drama would thus have to take into account several other models – from general hedonistic selection to drift, cumulation, or co-evolution, and possibly even cultural attraction (cf. Mesoudi, 2016; Morin, 2016; Sobchuk, 2023).

In general terms, the evolutionist template advanced by Moretti might be substituted with a more tempered vision of the process of change, based on ‘horizontal’ movements through the system, and emphasising the notion of variation over evolution, as in Küpper’s cultural net theory. In this perspective, the local branching of forms, while still present, would not necessarily be tied to geographical and linguistic spaces, but could also be explained by other poetological (e.g. the reception or rejection of Aristotle) or sociological factors (e.g. the material conditions of theatre).

Beyond these theoretical considerations, the reader should necessarily take into account the limitations of this research while assessing its contribution. The first one, as already mentioned, concerns the composition and the size of the EmDraCor corpus: while the theoretical and pragmatic choices informing its building process have been extensively discussed in Chapter Two, I am aware that a more accurate evaluation of evolutive dynamics would require retrieving a larger number of plays – and specifically of tragedies, if one wanted to restrict oneself to the assessment of Moretti’s original argument on ‘Baroque tragedies’.

More generally, a larger number of texts would reinforce the statistical reliability of my experiments and also allow one to test other modes of explanation for literary evolution (such as the ‘cohort succession’ hypothesis advanced by Underwood et al., 2022). An expansion to other linguistic domains beyond the five so far investigated would be desirable as well, especially if including traditions defined by transnational circulation, such as the Dutch one – situated at the crossroads between the French ‘regular’ and the English ‘irregular’ models.

From a methodological point of view, moreover, the operationalisation of the concept of drama I proposed in Chapter Three stands open to criticism. First, it could be argued that the specific choice of structural features I employed to represent the plays was inadequate to verify Moretti’s thesis, as it did not cover all elements he listed as

marking the divergence of early modern drama (such as those linked to the plays' language or performance). Secondly, it should be kept in mind that, while most scholars naturally focused on "the surface features of dramatic texts" because of their "easy computational accessibility", they are still "less reliable than they might seem at first glance" (Andresen and Reiter, 2024b, 2). A glaring example of this, which I previously discussed, is the comparability of segmentation practices across corpora and its effect on metric computation.

Considering the project's original conception, the first objection can be addressed on practical grounds: after all, the majority of the elements mentioned by Moretti are still form-related and have thus found concrete expression in the array of features employed. It does, however, underline once again an aspect the reader might want to keep in mind: the research behind this book focused exclusively on the written aspect of theatre, i.e. the dramatic text, and did not address questions related to its performative aspects or the social implications, which are central to other digital humanities investigations, such as Escobar Varela's (2021), and played an important role in early modern culture. This restriction of scope, however, is justified by the research object itself, since, as Lochert (2009, 29) notes, the European theatre of the sixteenth and seventeenth centuries is still characterized by a primacy of the text.

The second objection, regarding the broader validity of surface textual features, could instead be approached by pointing out the inherent limits of the practice of modelling. "Built into the modeling process", Richard Jean So (2017, 671) argues, "is a self-reflexive account of what the model has sought to measure and the limitations of its ability to produce such a measurement". Accordingly, the model itself should not be considered "as a potential antagonist, an entity that has brought forth intractable empirical truths that one must accept or reject", but rather a mediating instance which allows the scholar to confront data in a more efficient way, and which should be continuously updated and refined according to its performance. In this case, the operationalisation of dramatic form I proposed, and thus the formal features I selected to represent the plays, are justified by their usefulness in recognising patterns in

drama history, and can thus be used as a point of departure for further research.

Ultimately, the global significance of this work is twofold. On the one hand, considering the still unsatisfying data landscape for computational comparative drama analysis, the creation of EmDraCor – a multilingual, open-access, machine-actionable corpus of 150 TEI/XML-encoded plays – represents in itself a significant contribution to future digital research in early modern studies.

On the other hand, this project – like many others in digital humanities research – began by trying to provide an empirical reassessment (and, to some extent, a validation) of a previous qualitative theory. In doing so, it placed itself within the “triangulation” process outlined by Fotis Jannidis:

Characteristic of literary research is precisely the abundance of interpretative theses, especially on the canonical texts, an abundance of statements that contradict each other. So what does it mean when, in this context, such a thesis is taken up again through quantitative research and substantiated with quantitative arguments? In this way, the well-known thesis is supported anew, from a methodologically different perspective. In the social sciences, this phenomenon is referred to as ‘triangulation’ in the context of the discussion about ‘mixed methods’ and is seen by many advocates of a pragmatic approach to qualitative and quantitative methods as a gain in knowledge. (Jannidis, 2022, 10–11)

As shown in the previous pages, the experiments’ outcomes lent some support to Moretti’s global hypothesis, but also suggested that ‘horizontal’ variation could be a better interpretative key for the evolution of early modern drama than ‘vertical’ speciation – thus partly moving beyond the mere triangulation paradigm, and delivering some new insights. The book’s main contribution, however, amounted to the further development of a largely ignored method, i.e. the vectorisation of texts based on their structural properties. Accordingly, the main hermeneutical gain it provides is not found in the actual scrutiny of ‘Modern European Liter-

ature', but rather in the development and refinement of a methodology which might represent an efficient approach for dealing computationally with issues of literary form. This technique, I argue, might be particularly productive for comparative literary studies, insofar as it allows scholars to operationalise the fuzzy concepts of 'comparison' and 'similarity' with regard to formal aspects. Operations such as the comparison of vectors according to different implementations, or the inspection of clustering, might thus offer quantitative avenues for exploring humanistic models of comparison such as Wittgenstein's 'family resemblances' or Rosch's 'prototype theory'.

In this perspective, the work situates itself within what Anne-Sophie Bories and colleagues have called the "polite revolution of computational literary studies", i.e. "a moment of acute methodological creativity, with researchers branching out toward novel and often heterogeneous tools, experimenting with a number of possibilities, some of which, in turn, inform and sometimes change what they are striving to grasp" (Bories et al., 2022, 13). The methodological contribution this book offers should thus represent the most interesting takeaway for the reader, going beyond the results of specific experiments conducted; after all, "[t]he goal of experimental inquiry is not to get everything right the first time[,] but to advance knowledge in a way that invites further testing and refinement" (Underwood, 2019, 183).

Almost one hundred years ago, Boris Yarkho closed his highly experimental essay on speech distribution in five-act tragedies by bluntly stating: "[w]e have demonstrated the technique; the path has been cleared for those more fortunate than us" (Yarkho, 2019, 72). In a similar vein, but perhaps with a less laden rhetoric, I hope that other scholars will improve the methods I began prototyping here and fully explore the potential of formal vectorisation for the comparative study of literary form across traditions.