Experiments

The operationalisation of drama proposed in the previous chapter, consisting in the vectorisation of plays according to their formal features, aims at enabling meaningful and systematic comparison across diverse corpora. This serves two complementary goals – providing an empirical framework for testing Moretti's thesis of progressive diversification in European early modern drama while addressing the broader question of "how dramatic forms change and are preserved within a framework of cultural evolution" (Szemes and Vida, 2024, 18).

To reach these aims, different strategies can be employed. In my view, the simplest method would involve measuring the degree of similarity or dissimilarity between play embeddings, since it can be argued that two texts whose vectors are highly different from each other will also be different in terms of dramatic form. One implementation in this sense entails directly computing distances between coeval plays from different linguistic traditions, and then assessing whether such values are increasing across time. Alternatively, one could also use dimensionality reduction techniques to project the play embedding on a Cartesian plane and then look for clustering by comparing snapshots from consecutive time frames.

In the following pages, I experiment with both approaches, testing their hermeneutical power while pointing out their limitations, before eventually focusing on a more traditional analysis of the development of individual metrics. While exploratory by design, these experiments yield some first empirical insights into data, thus revealing whether the branching of dramatic traditions theorised by Moretti is mirrored

through a 'progressive distancing' of the plays' vectors, or what other dynamics are at play.

Distances

Data scientists and, by derivation, digital humanists and cultural analytics scholars have plenty of options at their disposal for measuring similarity or distance among any pair of vectors.[1] The choice of the right method, however, depends first and foremost on the properties of the vectors themselves. Since, as Nan Z. Da polemically argued, "it is good practice to ask users of CLS to show their vectors", because "it demystifies much of the process and often reveals conceptual weaknesses" (Da, 2019, 622), let us observe the composition of a sample vector from the EmDraCor corpus, representing Ruiz de Alarcón's play *La verdad sospechosa* (1619):

> V = {41, 3, 14, 11, 3, 0, 0, 1, 3306, 16919, 16461, 473, 6.571428571428571, 0.505494505494505, 0.7699649556792411, 14, 11, 1, 46, 0.04945054945 05494, 0.6608466660481829, 0.2431203018946239, 0.9268292682926 83, 0.5714285714285714, 0.6391666666666669, 3.0, 0.59740259740259 74, 0.8461538461538461, 0.3571428571428571, 0.2857142857142857, 0.3571428571428571, 4.99, 0.3571428571428571, 0.2857142857142857, 0.3571428571428571, 0.0175417232479634, 0.5493558858726955}

Within this 'bag-of-numbers', one immediately notices a sizeable difference in the values' magnitudes, ranging from small (0, 1, 2) to very large (16919), with many of them being long decimals. As statistical practice prescribes, I applied min-max scaling[2] to normalise them in the range

1 In their survey, Choi et al. (2010) list 76 such measures employed across different scientific fields.

2 Min-max scaling was preferred over *z*-score standardisation for three reasons. First, the play vectors combine heterogeneous metrics with different distributions (some skewed, some bounded), while *z*-score assumes each feature is roughly normally distributed. Second, rescaling to a fixed [0,1] range en-

from 0 to 1. At the same time, some play embeddings contained null values, indicating that not all features were present. However, since the number of zero values across all metrics in the corpus is relatively low (less than 5%), one can safely consider the EmDraCor vectors as substantially populated across all dimensions.

The vectors' characteristics, such as their relative denseness, guided the choice of the appropriate measure for computing the distances between them. Cosine distance was excluded because, by measuring only the angle between vectors in a scale-invariant fashion, it disregards magnitude differences that could be meaningful. Manhattan distance, for its part, is more suited to high-dimensional spaces, but the EmDraCor vectors only contain 38 dimensions; furthermore, it is less responsive to sharp differences on individual features, which are crucial for assessing the diversification hypothesis. Given the density, low dimensionality, and normalised scale of the play embeddings, I decided to adopt Euclidean distance[3] as the most appropriate metric, insofar as it requires no assumptions about distribution and remains reliable on this scale.

To test whether a general trend towards formal diversification could be recognised across the whole corpus, I first performed an experiment based on the repeated measurement of vector distances. To do so, I divided the 150 vector-plays into groups according to their normalised year; within each group, I calculated the Euclidean distances between each pair of plays. Then, I used the average of these distances to represent the 'formal difference' within each time frame, and plotted these averages to visually observe how they changed over time.

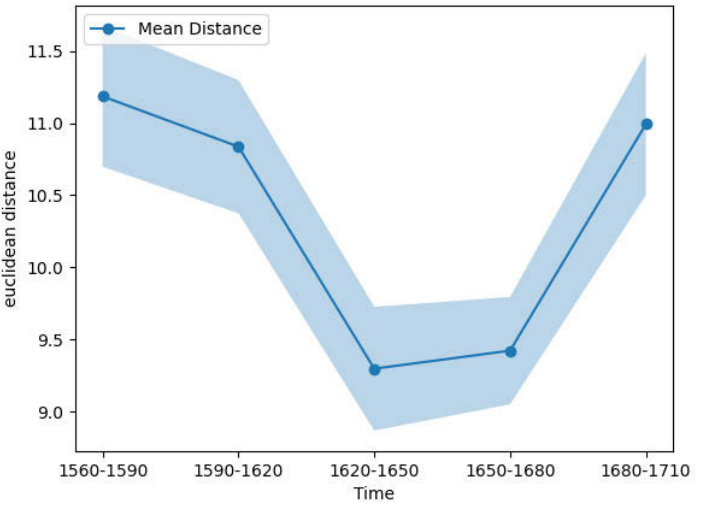
---

sures that no single feature dominates Euclidean distance calculations, while also yielding more interpretable values for a non-technical readership. Third, *z*-score standardisation tends to compress outlying values, which in this context may represent meaningful formal features rather than noise. On the other hand, however, min-max scaling is very sensitive to extreme values, as a single anomalous data point can compress the rest of the distribution.

3 This type of distance can be described as the length of a segment connecting two points on a Cartesian plane. It can be computed with the formula $d = \sqrt{(v2 - v1) \cdot (v2 - v1)}$ – where $v1$ and $v2$ represent any given vector (cf. Grootendorst, 2021).

Identifying the most suitable time sequences for visualising evolution would be worth another book in itself. While many DH studies use the century "as the typical yardstick of quantitative cultural history" (Moretti and Sobchuk, 2019, 90), there is no definitive answer to what Beausang (2022, 646) calls a "deceptively straightforward question", i.e. "at what scale does literature change?". In this context, and given the corpus' relatively short temporal arc, I employed fixed thirty-year-long time spans – a frame which corresponded to that of the 'generational cycles of literary evolution' postulated by Moretti himself in *Graphs, Maps, and Trees* (Moretti, 2007, 20–21).

Figure 5: Mean Euclidean distances between groups of coeval plays (30-year time frame, fixed windows), with confidence interval.



The results of this implementation, shown in Figure 5, describe a different narrative from that put forward in 'Modern European Literature'. Instead of a steady increase in diversification from the beginning of the seventeenth century, the data suggest a different trend, where a clear

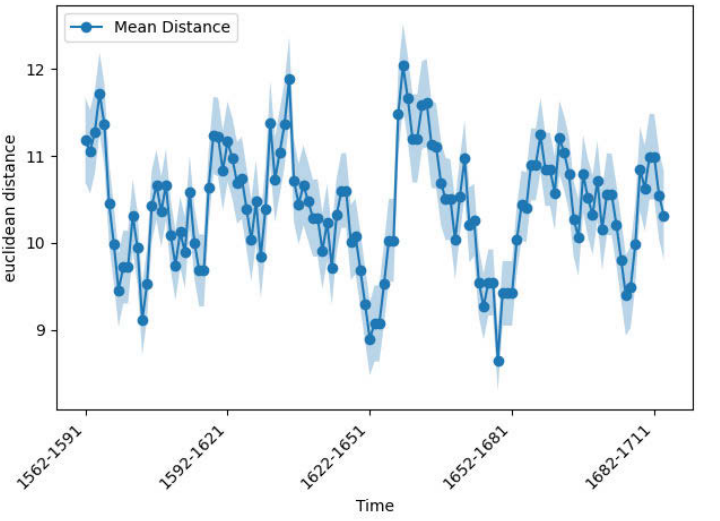
drop in formal differences in the first half of the century – a period of formal convergence, one might say – is followed by a sharp increase in diversity toward the beginning of the next one.[4]

Interpreting time series, however, is a notoriously difficult task, especially when they emerge from limited datasets. A different implementation, using instead a sliding rolling window of thirty years (Figure 6), returns a considerably more complex picture, where periods of formal convergence and divergence keep alternating. This greater volatility reflects both the sensitivity of the technique to small changes – with relatively few data points per window, the addition or removal of a single distance can noticeably shift the mean – and the finer temporal resolution it provides. A fixed frame analysis, based on larger and non-overlapping groups, trades granularity for stability, and smooths such oscillations into a single U-shaped pattern; however, as previously noted, a *t*-test shows that the downward and upward slopes are statistically significant. Furthermore, the same pattern still emerges (even if less forcefully) when the frame is scaled down to a 15-year window.

Computing pairwise distance between play embeddings within the same time frames, however, is not the only way of measuring distances inside the corpus. A complementary approach involves calculating an average vector for each group of coeval works from the same linguistic area and considering it a proxy for the 'position' of a given dramatic culture within a period. Then, fluctuations in the distance between such 'centroids' can be measured to assess diachronic divergence in formal features. This approach is not unprecedented in computational drama analysis: Ilsemann (1995, 17) followed an analogous path in "establish[ing], if not invent[ing], the shape of the average play on the basis of [values from] the plays already examined", so that "deviations from the model could be made visible". By collapsing plays into averages, however, this method trades individual variation for a clearer view of aggregate trajectories.

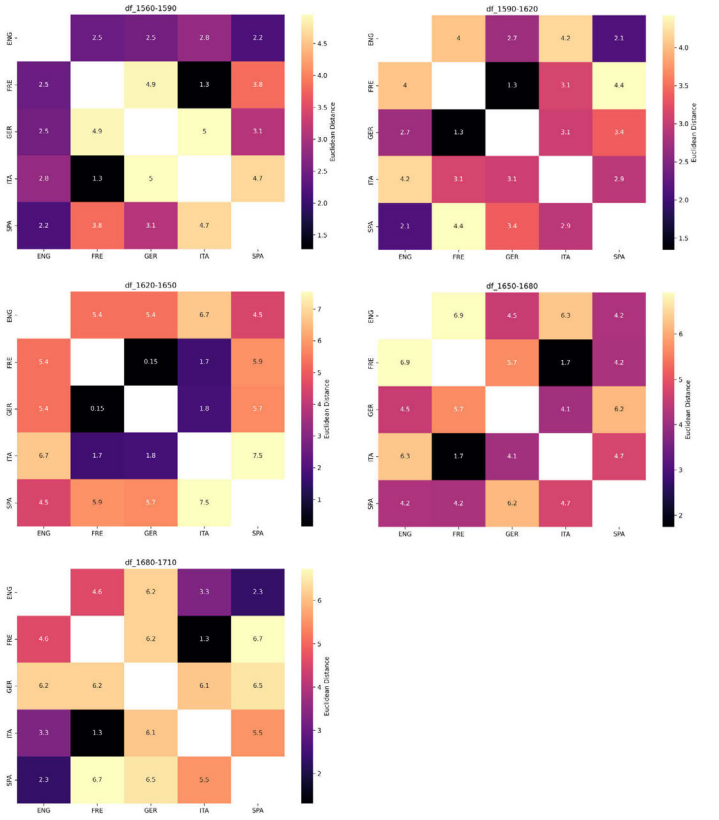
4 The reliability of the two slopes (but not of the plateaus) is confirmed by Student's *t*-test (which determines whether the difference between the means of two groups is statistically significant) and visually displayed by the relatively small confidence intervals around them.

Figure 6: Mean Euclidean distances between groups of coeval plays (30-year time frame, rolling window), with confidence interval.



If Moretti's hypothesis holds, one would expect centroid distances between traditions to increase over time, and the heat maps presented in Figure 7 partially support this interpretation. Unlike the pairwise analysis, which revealed a U-shaped pattern, centroid distances suggest a more consistent divergence. Looking at the matrices, a general shift from darker to lighter tones, reflecting a movement from similarity to dissimilarity, can be noticed. At the same time, the graphs highlight known relations between theatrical cultures, such as the continued formal closeness between the Italian and French traditions (the 'regular' theatres in Lochert's reconstruction).

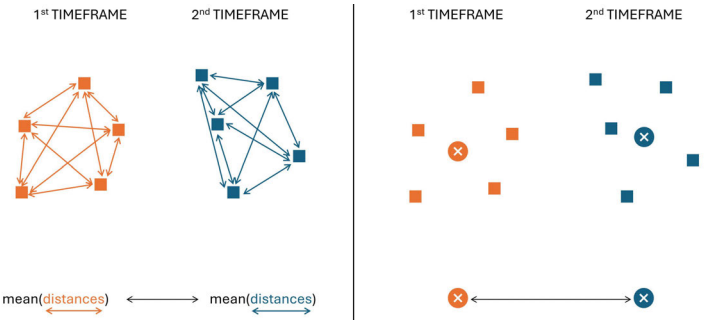
Figure 7: Heat maps showing distances between 'centroids' for different linguistic groups at different chronological stages (30-year time frames).



The contrasting results from the previous experiments can be explained by considering the different logic underlying each, which are summarised in Figure 8. In the first case, the iterative computation of pairwise distance between any two vectors returns a list of distances, whose mean is then compared across time frames. Being based on the repeated comparison of individual works, this procedure could be seen

as more granular but also more sensitive to outliers. In the second case, one identifies average values for groups, creating synthetic vectors for any spatio-temporal context, and then measures their distance. Each method captures a different aspect of the same data, but neither alone is decisive in supporting or refuting Moretti's claims, suggesting the need for another approach.

Figure 8: Visualising the logic behind two options for computing vector distances: pairwise (left) and centroid-based (right). Each dot represents a play embedding.



Clusters

As shown, preliminary results from the exploration of distances between plays paint a more complex picture of drama evolution than the progressive diversification postulated by Moretti. To examine the question from another perspective, I decided to switch to a graphical representation of the data, plotting the vectors on a Cartesian plane and tracing their clustering in successive temporal snapshots. The reasoning behind this strategy was straightforward: if a diversification of European drama had indeed taken place across the early modern era, this should be signalled by the progressive emergence of clusters organised by linguistic tradition.

To plot multi-feature vectors on a two-dimensional space, however, one first needs to apply an appropriate algorithm for scaling down the data. Machine learning research has developed several algorithms for dimensionality reduction (cf. Ayesha et al., 2020; Van der Maaten et al., 2009), which can be classified according to various criteria.[5] More generally, an overarching distinction is often formulated between methods trying to preserve the general structure of the data and those focusing on preserving its local geometry. While any compression of high-dimensional vectors invariably leads to a loss of informational accuracy, it also allows for an intuitive, visual grasp of how data points relate to each other.

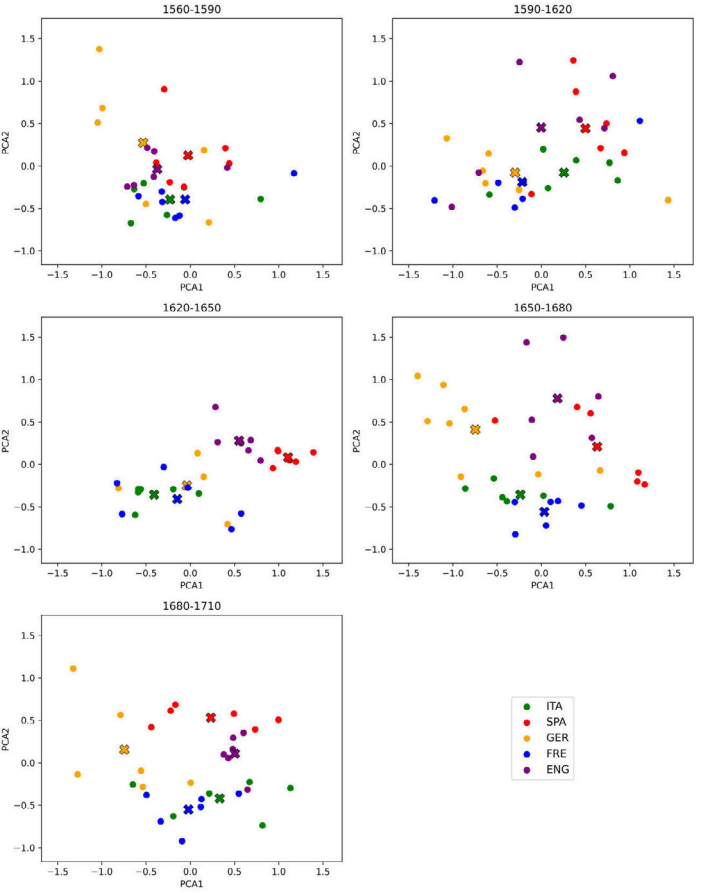
In the context of computational literary studies, Principal Components Analysis (PCA, see Lever et al., 2017) is often employed as the standard technique for dimension reduction. This preference likely stems from its early adoption in the subfield of computational stylistics, as testified already in work by John Burrows and others (Beausang, 2020, 185 ff.), from which it later entered the CLS mainstream. The relative ease in implementing PCA makes it arguably the most popular option for scholars looking to build "a low-dimensional representation of the data that describes as much of the variance in the data as possible" (Van der Maaten et al., 2009, 3), even though issues in its correct application by DH scholars have been noted (see Leufkens, 2020).

Despite also running some tests with other unsupervised techniques such as t-SNE (van der Maaten and Hinton, 2008) and UMAP (McInnes et al., 2018), I ultimately chose to employ PCA not only because of its straightforward implementation and interpretability of results, but also because of its better performance in preserving global structures – and thus unveiling patterns of divergence or convergence. In my implementation, I used 30-year time frames as observation windows, thus plotting five different snapshots from my corpus: 1560–90, 1590–1620,

5 Anowar et al. (2021) individuates three main criteria: the linearity of the scaling process (linear/non-linear), the existence of predetermined class labels (supervised/non-supervised), and the type of mapping (random projection-/manifold-based).

1620–1650, 1650–1680, 1680–1710. Data points were marked according to the dramatic tradition they belong to, while centroids for each group were added as well.

Figure 9: EmDraCor play embeddings projected on a Cartesian plane through PCA and grouped into 30-year timeframes, with centroids.



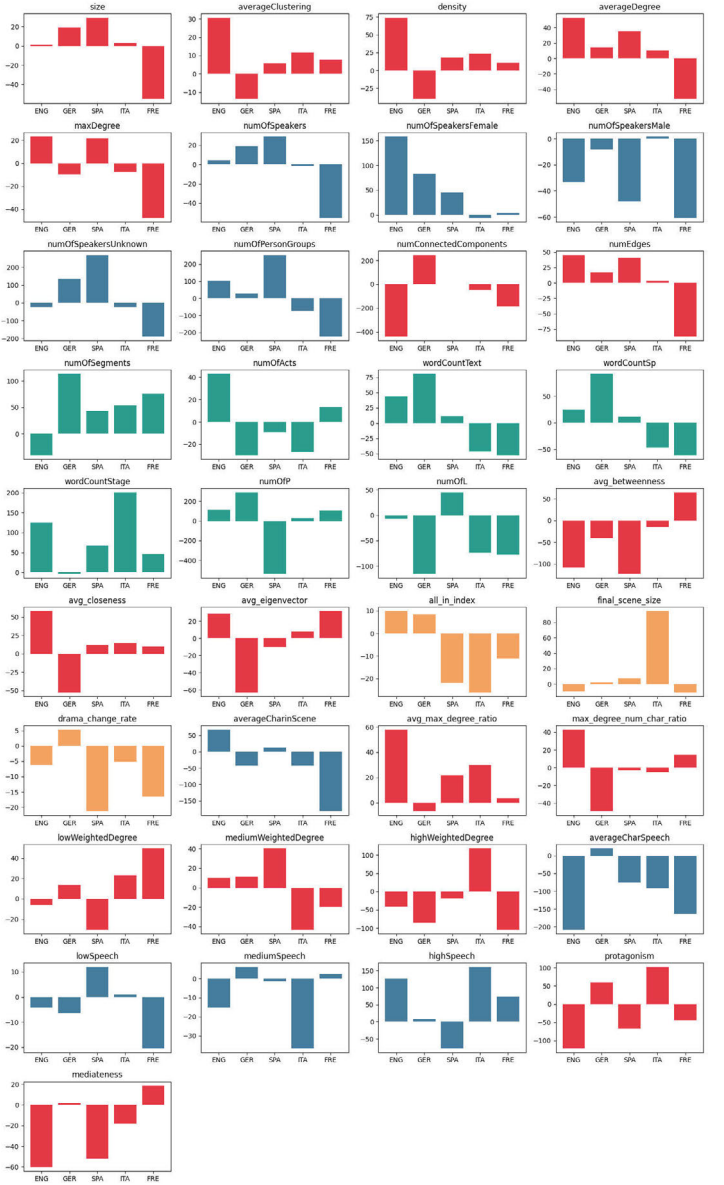
Results presented in Figure 9 partially support, once again, a narrative of diversification. While the extent of the vectors' dispersion across the PCA space remains relatively stable in the five plots, rough clusters of plays begin to appear, especially from 1650 onwards. This stands out especially when comparing the first graph, where plays from all regions cluster in the central area, and the last one, where the zones occupied by different dramatic traditions almost do not overlap – with the notable exception of French and Italian, whose formal proximity, shaped by a common adherence to neo-Aristotelian poetics, has already been noted in the heat maps.

Furthermore, inspecting the feature loadings of the two components plotted (which together account for almost 50% of the variance) shows which dimensions mostly contribute to positioning the play embeddings across the PCA space. While the first one (PCA1) loads mostly on network *density* features, such as density (19.2%), closeness (13.2%), and the ratio between average and maximum degree (13.7%), the second one (PCA2) relies on network *size* features, such as number of edges (11.2%), average of characters per scene (11.5%), and maximum degree (10.2%). This suggests, with some approximation, that the two main axes of formal variation within the corpus are network concentration and network scale.

Patterns

The evidence presented so far captures some variation of early modern drama at a formal level, but it still does not reveal a trend as clear as the one hypothesised in 'Modern European Literature'. While vector distances calculated using different approaches (pairwise or centroid-based) tell divergent stories, examining the clustering of the play embeddings via PCA seems to confirm a tendency towards diversification, but the pattern is not pronounced enough to be considered conclusive.

Figure 10: Percentual variation (from 1561 to 1710) for the value of each vector component across the five linguistic traditions.



Another approach, however, is possible. Instead of performing operations on the vectors as a whole, one could decompose them again into their constituent metrics and then examine their individual diachronic evolution. In other words, even if it is not possible to pinpoint a global trajectory for the whole corpus, it might be possible to single out which features are distinctive of a given literary tradition, i.e. which formal aspects are most decisive in distinguishing it from the other coeval ones.

An intuitive approach would involve plotting the values for a given metric after arranging the texts in each subcorpus in chronological order. Applying this method systematically, one can then compute shifts in absolute value for each metric across the time span investigated, and thus verify whether the variation of a given feature has actually become highly distinctive of a national tradition against the others. While I could have calculated these shifts directly, measuring the change between each pair of adjacent values, a few metrics had null values, thus creating spurious slopes.[6] Therefore, I fitted a linear regression to each metric's values over time and plotted the resulting rate of change, expressed as a percentage, for each tradition (Figure 10).

The risks of such a simplifying approach are well known and have been eloquently outlined by Moretti and Sobchuk (2019): trends obscure the complexity of data insofar as they "make data easy to read" and "give them a meaning", while reducing "multiple trajectories to a single big story" (93–94). Even when trends are "unquestionably there", and "quite uniform and gradual […] this doesn't necessarily mean that the forces behind them are equally uniform" (97).

As far as the EmDraCor corpus is concerned, the best antidote to any overinterpretation of trendlines lies in a thorough analysis of the data behind any of them. Some slopes, for example, appear skewed by the presence of outliers: the values for mediatedness in the French subset remain mostly stable during the 1561–1710 time frame, but a couple of

6 E.g. a play without any stage direction, and thus with 'wordCountStage' = 0.0. This happens in about 6% of the overall values.

peculiar early works[7] creates the appearance of a slight upward trend; a similar argument could be made also for the plot-related measure 'all-in index' in the Italian subcorpus.

Similarly, the distribution of the texts in terms of genre can lead to some red herrings. The Italian subcorpus, for example, seemingly displays a notable growth in the final scene size values, but by observing its composition one notices how this is likely due to the random concentration of comedies (known to feature large ensemble casts in the last scenes) in the later decades.

Instances like this suggest that each percent variation needs to be double-checked with the actual trend lines and the corpus composition before it can be used to support broader arguments about literary history. Furthermore, not all the shifts in the features' evolution can be taken as equally significant: changes in the number of prose lines ('numOfP') and verse lines ('numOfL'), for example, are not very telling, since the two modes are most often mutually exclusive (i.e. a play would usually be either fully in prose or fully in verse).

While keeping these caveats in mind, however, I believe tracking the evolution of individual metrics across EmDraCor still represents a useful first step towards a quantitative profiling of dramatic traditions. Such findings are preliminary, constrained by the size and composition of the corpus; once the data landscape improves and bigger multilingual corpora become available, it will be possible to re-run these experiments on them, following standard reproducibility practices. In the meantime, the following pages collect and briefly discuss relevant features marking the autonomous development of each tradition – in a necessarily rhapsodical manner, moving freely across the most relevant metrics as the data warrants.

7 Both Bienvenu's *Comédie du monde malade et mal pansé* (1568) and Virey's *La Machabée* (1596) have fully-connected character networks, where every node is equally central to all connections, and because every shortest path is direct, no node serves as an intermediary on any shortest path. Thus, each node has a betweenness centrality of 0, and this translates into a null mediatedness value.

English drama

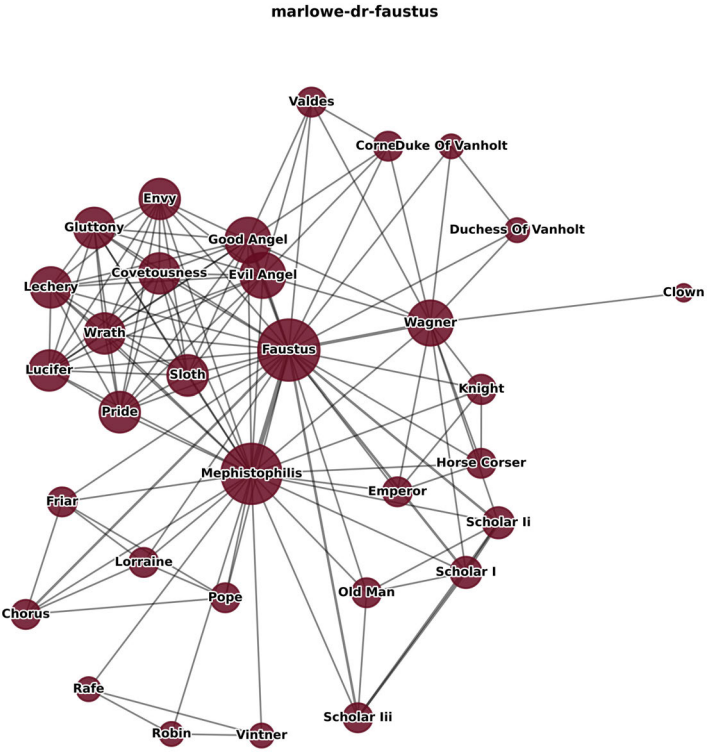
The English tradition displays the most significant metric variations across EmDraCor, suggesting that its theatrical culture was perhaps the one to develop the strongest formal autonomy during the time frame considered. This plays quite aptly into the narrative of the 'exceptionalism' of English drama, which has often been tied to the country's peculiar geography:

> Would there be Shakespeare, had England not been an island? Who knows? But that the greatest novelties of tragic form should arise away from the mainland, and from someone with 'small Latin and less Greek', is quite a sign of what European literature had to gain from losing its unity, and forgetting its past. (Moretti, 2013a, 13)

In terms of network structures, one notices immediately the parallel spike in average clustering and density, suggesting an enhanced connectivity both at a local (clusters) and general (network) level. The tightening of relations is also confirmed by the fact that the plays' size remains broadly unchanged. Furthermore, in keeping with the findings of Algee-Hewitt (2017), protagonism and mediatedness also drop across the time span.

Such shifts in network structure can be easily grasped by comparing play networks from the two chronological extremes of the corpus. Marlowe's *Doctor Faustus* (1588, Figure 11) is a typical example of diffuse Elizabethan tragedy, where the central character ties together different groups that would otherwise not interact. In this case, the eponymous scholar – embedded in a high-density clique with his demonic aide Mephistophilis and his servant Wagner – acts as the bridge between clearly distinct subnetworks, such as the courts of the Pope and the Holy Roman Emperor, the Vanholt ducal household, or the magicians Cornelius and Valdes.

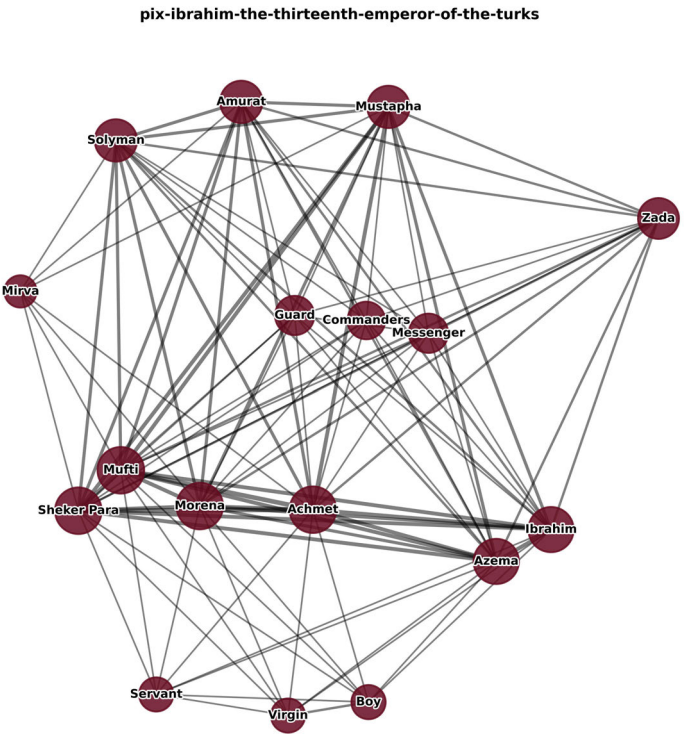
Figure 11: Co-presence character network of Marlowe's Doctor Faustus.



More than one century later, Pix's *Ibrahim* (1696, Figure 12) shows instead the 'distributed' features whose increase Algee-Hewitt tracked in this essay. While no single character stands out as a towering protagonist, a closely-knit group of central figures assumes such function. These include the titular character (the sultan Ibrahim), the villains (Sheker Para, Azema, and Achmet), their opponents (Solyman, Mustapha, and

Amurat), and the heroine Morena, whose scholarly-validated centrality in the plot economy[8] is not fully reflected by her network position.

Figure 12: Co-presence character network of Pix's Ibrahim.



8 As Lechner (2016) argues, the treatment of the character of Morena (first violated by Ibrahim at the instigation of Sheker, then inciting a rebellion against the sultan, and ultimately emulating Lucretia in killing herself) signals the transition from the traditional heroic drama *à la* Dryden to the pathetic *she-tragedy* of the late seventeenth and early eighteenth century.

Furthermore, cast composition shifts significantly across the corpus' time frame: while the total number of characters remains roughly the same, the quota of females increases by a notable 80%, while that of males drops by 10%. This trend is strongly influenced by the plays of 1660–1680, where the number of female characters increases rapidly, peaking in 1675 with Thomas Shadwell's graphic rewriting of the *Don Juan* myth, *The Libertine*. In this play, the stage is indeed almost always occupied by some of the rake's female victims, who are brought forth both in groups (e.g. the nuns and shepherdesses he attempts to assault, or the "whole Batalion of couragious Women" – to quote his servant Jacomo – he tricked into believing they were his wives) and as individuals (Leonora, with her lady-in-waiting Maria and maid Flora; the sisters Clara and Flavia).

Such an increase can likely be explained by the groundbreaking introduction of actresses on stage starting in the 1660s (see Howe, 2001, ch. 1). After all, the choice of the *dramatis personae* was determined not only by poetics, but also by pragmatic reasons: since "most playwrights wrote their plays with a particular company in mind and fashioned roles to fit the specialties in that company" (Howe, 2001, 11), the sudden availability of female performers might have prompted playwrights to insert more female roles in the pieces.

At the same time, the little drop in the overall drama change rate suggests that only small cast changes kept taking place between scenes. One may speculate that this trend reflects the growing, but not uncontested influence of neoclassical French theatre, where the practice of the *liaison des scènes* required at least one character to always remain on stage and link different narrative segments. As noted already by Friedland (1911), the attempt to keep 'scenes unbroken' – alongside with the adoption of the three main unities of time, space, and action – was an important component in the reception of French theatrical models and was explicitly discussed by dramatists who tried to adhere – not without reservation – to such prescriptions.[9]

9 John Dryden, in the first prologue to *Secret Love; or, the Maiden Queen* (1688), assures he followed "the exactest rule, by which a play is wrought", including "the

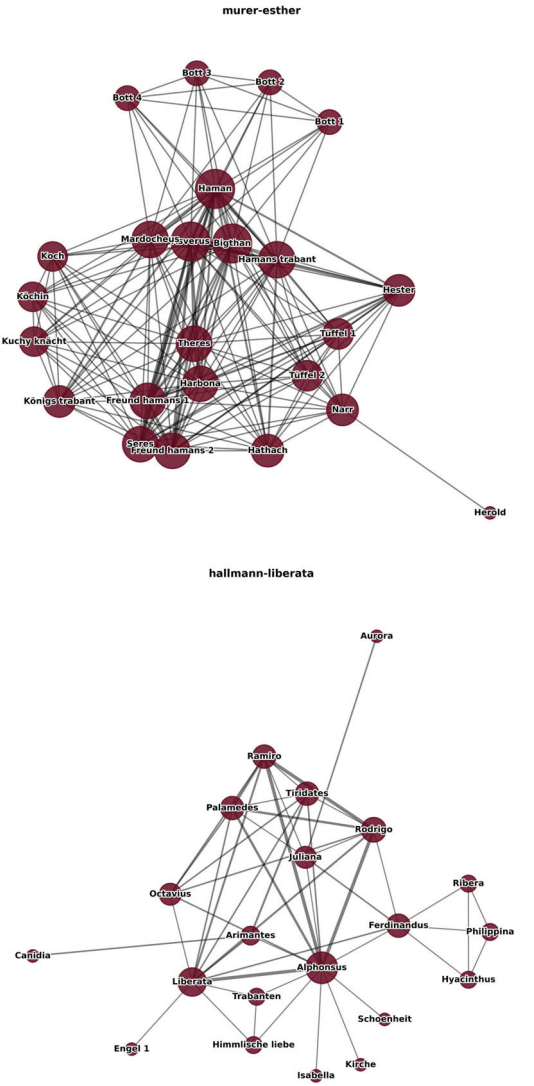
German drama

By contrast, German plays gradually shift towards sparser networks, as slight size increases are accompanied by decreases in density and average clustering — characteristics unique to this subcorpus. Other indicators of this trend include a rise in the number of independent subnetworks, which might also explain a reduction in centrality measures such as betweenness, closeness, and eigenvector centrality. A visual comparison between a play from the late sixteenth century and one from a century later may be useful to illustrate this point; Figure 13 contrasts Jos Murer's *Esther* (1567) and Johann Christian Hallmann's *Liberata* (1699).

While the former network features a compact structure, where high density likely reflects the intricate power dynamics within Ahasuerus' court, the latter exhibits a star-shaped pattern, where a core group of characters largely acts as the sole link between peripheral individuals. Furthermore, density is further reduced by the presence of individual characters who occupy distinct segments by themselves, frequently interacting with the real-world audience through humorous asides (e.g. by Pollio Asinius) or moralising tirades (e.g. by the personified Justice).

scenes unbroken", but also contrasts these arid prescriptions ("dead colours") to his own contribution, in terms of "wit" and "plot", which he characterised as "the living beauties of a play". By his part, Thomas Shadwell prefaces *The Sullen Lovers: or, The Impertinents* (1688) by saying he has "as often as [he] could naturally, kept the Scenes unbroken, which (though it be not so much practised, or so well understood, by the English) yet among the French-Poets is accompted a great Beauty". He later warns, however, "not to expect a very Correct Play", since "[i]t is so difficult a thing to write well in this kind". Both examples are mentioned by Friedland (1911, 286; 291).

Figure 13: Co-presence character networks of Murer's Esther (above) and Hallmann's Liberata (below).



At the same time, works become longer, mostly through an increase in lines spoken by characters, while the share of stage directions remains constant. This fact, however, cannot alone explain the strong rise in the number of segments, i.e. scenes, which could instead be attributed to the influence of French theatrical poetics. Following neoclassical precepts, it seems indeed that playwrights were prompted to create a new scene each time the character configuration changed, thus significantly increasing the number of scenes.[10] This tendency reaches its apex in authors such as Christian Weise, whose *Ungleich und gleich gepaarte Liebes-Alliance* [*Unequally and equally matched love-alliance*] (1708) features 53 segments. Reading the play, one notices how every slight variation in character configuration comes to constitute a new *Auftritt*, with most scenes featuring only a couple of characters engaging in short, stichomythic exchanges.

Spanish drama

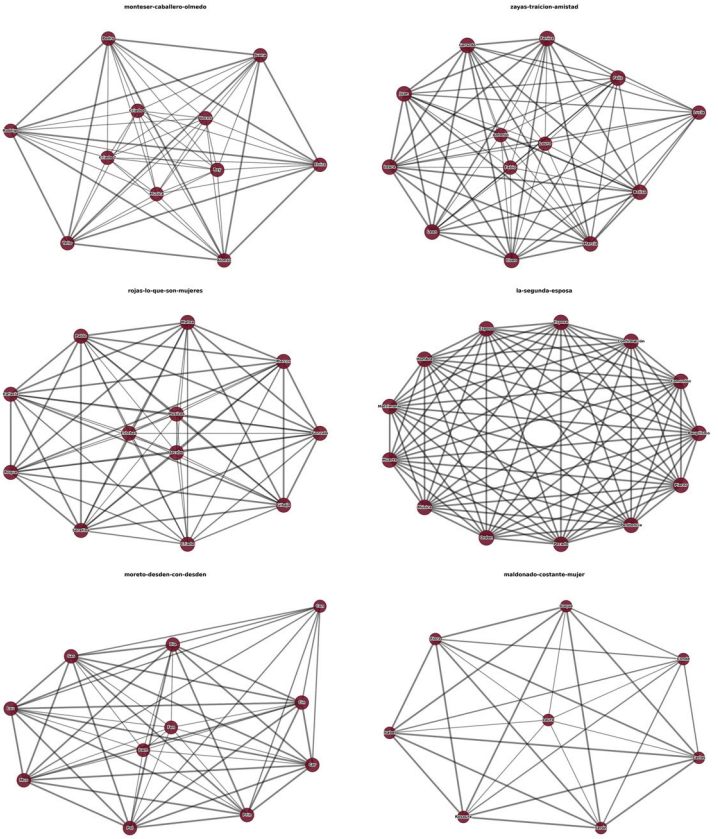
The Spanish subcorpus displays growth in all basic network measures (size, average clustering, density, average and maximum degree), indicating expanding and more densely connected dramatic networks. Furthermore, it shows an interesting pattern in terms of centrality values, with a drop in betweenness centrality and mediatedness. This is clearly visible in the structure of many Spanish character systems, especially from the 1620–1660 period, which take the shape of stable, often fully connected networks (Figure 14). In these examples, no character seems to act as a clear 'mediator' among groups; conversely, the different characters associate freely, forming complex constellations.

The absence of a central mediating figure is also underscored by the distribution of weighted degree as classed by Szemes and Vida (2024). While the number of characters with high and low weighted degree tends to diminish, characters in the middle tertile, i.e. having a

10 For a detailed reconstruction of the development of the concept of 'scene' in German drama and its related terminology (from *scena* to *Teil*, *Aufzug*, and finally *Auftritt*) see Jordan (1939, §2–3).

moderate number of connections, increase, pointing to networks where character centrality is more equally distributed.

Figure 14: Network structures of six 17th century Spanish plays. From the upper left: Monteser's El caballero de Olmedo (1621), Zayas' La traición de la amistad (1630), Rojas Zorrilla's Lo que son mujeres (1644), Calderón's La segunda esposa (1648), Moreto's El desdén con el desdén (1654), and Maldonado's, Dueña's and Cifuentes' La más constante mujer (1658).



Another factor setting the Spanish corpus apart from the others is the massive increase in ungendered speakers ('numOfSpeakersUnknown'). According to the corpus' encoding conventions, the 'unknown' tag marks two different categories: abstract or allegorical entities without a clear gender (e.g. 'Virtue'[11]) and individuals/groups whose gender is not possible to determine. The first type features prominently, for example, in Juan Montenegro y Neira's *La toma de Buda* [*The Conquest of Buda*] (1687), where the stage is populated both by historical characters (Turks and Christians involved in the eponymous siege) and allegorical figures like Unity, Discord, and the Seven Deadly Vices.

The second case is exemplified by the many unspecified voices that populate plays like those by Antonio de Zamora, often characterised as collective speakers – thus contributing to the notable increase in the 'numOfSpeakersGroups' parameter as well. In his two works, *Ser fino y no parecerlo* [*Being refined and not looking like it*] (1701) and *Cada uno es linaje aparte y los Mazas de Aragón* [*Each one comes from a separate lineage, and the maces of Aragon*] (1708), Zamora makes ample use of non-specified individual or plural voices for dramatic effect, employing them to better convey the hectic atmosphere of battles and sieges.[12]

Finally, the decrease in the drama change rate, while moderate, could be explained by a progressive implementation of the *liaison des scènes*, as seen in the English corpus as well. However, one does not necessarily need to explain such practice through the exogenous influence of French classicism. Lope de Vega himself was indeed a proponent of this compositional strategy, as he recommended that "[v]ery seldom should the stage remain without someone speaking, because the crowd becomes restless in these intervals and the story spins itself out at great length;

11 While allegorical figures in early modern texts are often depicted under feminine appearance (see Warner, 1985), they still embody abstract concepts, and thus their gender has been consistently annotated throughout EmDraCor as 'unknown', unless it was culturally acquired (e.g. the Devil as a male character).

12 See, e.g., the third act of *Ser fino y no parecerlo*, where several unnamed Phoenician and Achaean warriors clash.

for, besides its being a great defect, the avoidance of it increases grace and artifice" (Vega, 1914; cf. Ehrlicher et al., 2020, 7).

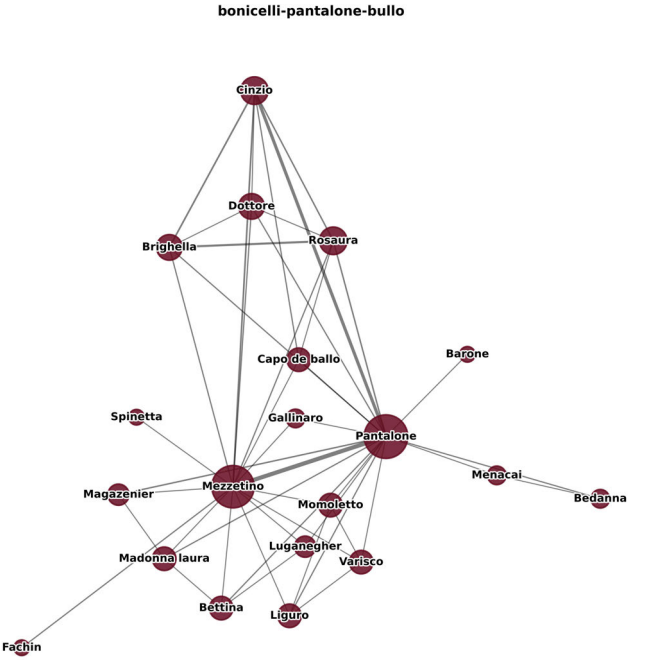
Italian drama

The Italian subcorpus, for its part, displays mostly moderate shifts in terms of network values, and generally exhibits less distinctive features compared to the others. Towards the end of the corpus, however, two metrics in particular show a sharp increase: the number of characters with a high weighted degree, i.e. with sustained interactions with other characters, and protagonism. Together, they suggest a pattern of progressive concentration of the protagonist role in one or few characters, in contrast with Algee-Hewitt (2017)'s findings on English early modern drama. Speech statistics support this reading, with an increase in the number of prolific speakers (reflected by the 'highSpeech' metric) at the expense of quieter figures.

A good example of this tendency is Giovanni Bonicelli's comedy, *Pantalone Bullo* [*Pantalone, the Bully*] (1688), where the relevance of the titular character – a vibrant rewriting of the stock type from the *commedia dell'arte* repertoire – cannot be overstated. Roaming across Venice with his gangs of thugs and causing mischief in taverns and shops, Pantalone is constantly in the foreground of the action together with his sidekick Mezzettino, as the network (Figure 15) shows. By contrast, the conventional, trite love subplot between Pantalone's daughter Rosaura and Cinzio is clearly confined to the periphery of the graph, and their interactions never become relevant in the network's overall economy.

Another interesting feature of Italian drama is the notable increase, especially towards the end of the seventeenth century, of stage directions. This might be explained with the presence of several works which, while not strictly belonging into the improvisation-based *commedia dell'arte*, still share its emphasis on the corporeal and performative component, and clearly foreground it in the written text. A clear example in this sense is Niccolò Amenta's *La somiglianza* [*The Resemblance*] (1706), where the author continuously includes stage directions (almost 200!) on how characters should come forth or interact with each other.

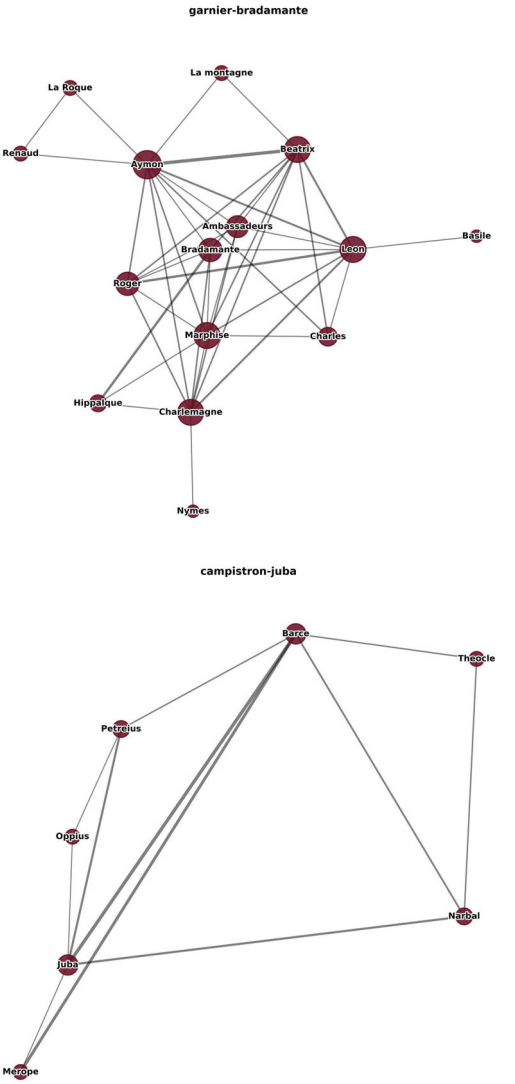
Figure 15: Co-presence character networks of Bonicelli's Pantalone bullo.



French drama

The most notable pattern found in the French subcorpus is a significant decrease in cast size, which does not affect all genders equally: while the percentage of female characters remains quite stable, the share of male and group characters decreases more sharply. This reduction produces leaner networks where, with fewer characters sharing the stage, all basic centrality measures (betweenness, closeness, eigenvector) rise mechanically: each figure is closer to all others, mediates more interactions, and is more directly connected to other central characters – even as the absolute number of connections and the number of edges decrease by virtue of the smaller casts.

Figure 16: Co-presence character networks of Garnier's Bradamante (above) and Campistron's Juba (below).



Although this trend can partly be explained by the presence of several one-act plays in the final decades of the corpus, it also reflects the extent to which the poetics of *classicisme* had progressively become ingrained in contemporary dramatic production, leading to more densely interconnected network structures. This shift can be seen clearly when comparing, for instance, the character network of an early play, such as Robert Garnier's *Bradamante* (1582), with that of the later tragedy by Jean-Galbert de Campistron, *Juba, Roi de Mauritanie* [*Juba, King of Mauritania*] (1685), as shown in Figure 16.

The evolution of French dramatic structures towards streamlined network models also affects the properties of the individual characters: since the cast is reduced, the proportion of prolific speakers increases (as testified by the increase in the 'highSpeech' value). At the same time, the decreasing drama change rate, suggesting more stability in stage configurations, is clearly linked to the authors' adherence to the *liaison des scènes* practice. The two phenomena can be observed again in the play by Campistron, a disciple of Racine, where all characters (and not only the central pair, represented by the tragic figures of King Juba and his spouse Barcé) deliver substantial speeches and serve as connectors between successive scenes, maintaining the *liaison*.

Coda: A reproduction experiment

Starting from the early 2000s, the debate on the 'replication crisis' in the sciences (see e.g. Ioannidis, 2005; Open Science Collaboration, 2015) has also reached other data-driven fields, including the digital humanities, where it prompted sustained reflection on the reliability of computational findings (Peels and Bouter, 2018; Herrmann et al., 2023; Schöch, 2023; Joyeux-Prunel, 2024). Within computational literary studies, a milestone in the discussion was Nancy Da's polemical attempt to build 'a computational case' against the entire subfield on the grounds that some results from influential studies could not be replicated (Da, 2019). Despite its significant conceptual and methodological shortcom-

ings,[13] Da's intervention was instrumental in consolidating the view that replicability and reproducibility cannot be taken for granted and need instead to be built into any sound study design.

Following this principle, all experiments presented in the previous pages are fully replicable through the data (TEI-XML files and CSV metadata) and code (Python scripts) provided in the *GitHub* repository. From a cultural analytics-oriented point of view, however, these options may seem insufficient: even if the rationale for the corpus building has been described in detail, one is still left wanting to re-implement the same methods on broader text collections to reinforce its empirical results.

The scenario I am describing here could be defined, according to the taxonomy developed by Schöch (2023), as a 'reproduction (of results)': "research question and method of analysis remain identical", but "the dataset, instead of also being identical, [is] more or less similar, but not unrelated" (389). The goal of such a procedure would be to assess whether findings obtained hold true across more than one dataset, and thus are valid more generally, and whether methods developed are robust enough for further reuse.

In my case, performing a reproduction experiment is a necessary step towards assessing the trustworthiness of my findings. While gaining insights into the EmDraCor corpus, following the practice of 'scalable reading', is a valuable goal in itself, it is also necessary to assess their significance in a broader context. Accordingly, a reproduction experiment would allow one to measure how many of the results presented in the previous pages are due to the idiosyncrasies of the text selection, and how many are actually generalisable and can be treated as reliable contributions towards a quantitative history of early modern drama.

Unfortunately, most of the plays featured in EmDraCor have not been sampled from larger machine-actionable collections; to a large extent, the collection itself has been assembled to address the lack of extensive

13 See e.g. the responses in the *Critical Inquiry* online forum (https://critinq.wordpress.com/2019/03/31/computational-literary-studies-a-critical-inquiry-online-forum) and the contributions by Jannidis (2020) and Piper (2020).

early modern 'programmable corpora'. Nonetheless, even if the experiment cannot be fully reproduced on wider datasets, it is still possible to do it partially by leveraging the only two corresponding DraCor corpora that, at the time of writing, contained a significant number of early modern texts, i.e. the French (FreDraCor) and English (EngDraCor) ones.

Together, these two corpora contain more than one thousand texts (602 French, 529 English) from the 1561–1710 time frame, and thus represent a significant benchmark collection for reproducing the experiments previously presented. Furthermore, deepening the analysis of these two dramatic traditions also holds some inherent theoretical interest, since French and English drama can be considered to embody, in Lotmanian terms, the two main configurations of European drama of the age, i.e. a 'regular' (French) and an 'irregular' (English) model. A reproduction experiment would thus enable further exploration of this dichotomy from a quantitative-formalist angle.

In practical terms, the procedure for obtaining the 'play embeddings' for French and English texts was the same as that developed for EmDraCor: I downloaded the metadata tables and enriched them through the usual Python scripts, excluding all texts outside the chronological boundaries of the early modern period as previously defined (1561–1710). Once the vectors were constructed, I reran all scripts to investigate again the three dimensions of interest (distances, clusters, patterns), and compared the results with those obtained from EmDraCor.

Before presenting the findings, however, a caveat is in order. The data available from the English and French drama corpora at the time of writing mostly overlapped only between the 1620s and the 1650s, preventing one from tracking the parallel evolution of these two dramatic traditions in a more comprehensive way. Furthermore, in the case of English drama, data availability meant that I could only compare the English subset of EmDraCor with early works from the main corpus (up to ca. 1650).

One should therefore tread lightly in analysing the results presented in Figure 17, which compares the evolution in the values of the 38 vector metrics across the four collections under investigation (EngDraCor, EmDraCor-Eng, FreDraCor, EmDraCor-Fre). Once again, some trends can

be misleading and require scalable readings in order to be appropriately contextualised, as two examples from the French tradition illustrate.

To begin with, the increase (ca. +25%) of the metric 'size' in FreDraCor seems to contradict not only the results from the French subset in EmDraCor but also the scholarly intuition about the shrinking of casts in the *théâtre classique*. A closer inspection of the actual values for each play reveals, however, that this last observation still holds for most tragedies, while the outliers pushing the trend upwards are comedies and hybrid forms featuring music. This last category was excluded from the EmDraCor sample and includes, for example, many *tragédies lyriques* by Philippe Quinault, such as *Isis* (1677) – the period's largest play with a total of 63 speaking characters.

In any case, the increase in the number of characters in FreDraCor represents the biggest departure from EmDraCor-Fre, and has a ripple effect on all network-related metrics, which therefore show slight growth. Data scarcity explains instead the other notable counter-trend, i.e. the evolution of the 'protagonism' metric: an initially higher value, determined by only about 15 works from 1580 to 1630, starts dropping as soon as the corpus size dramatically increases.

Nonetheless, many results from the larger corpora confirm the trends seen in their smaller counterparts. For example, all four collections show a noticeable growth of the number of stage directions ('wordCountStage'). While it seems premature to interpret such a trend as an early sign of the 'tendency towards epification' that Trilcke et al. (2020) recognised in later German drama, it still points to an interesting development: the fact that the playwrights started providing more and more cues to the performers suggests indeed a gradual shift in the relationship between written text and stage practice.

Figure 17: Evolution of vector components in the English and French subsets of EmDraCor (lighter tones) and the respective main DraCor corpora (darker tones).

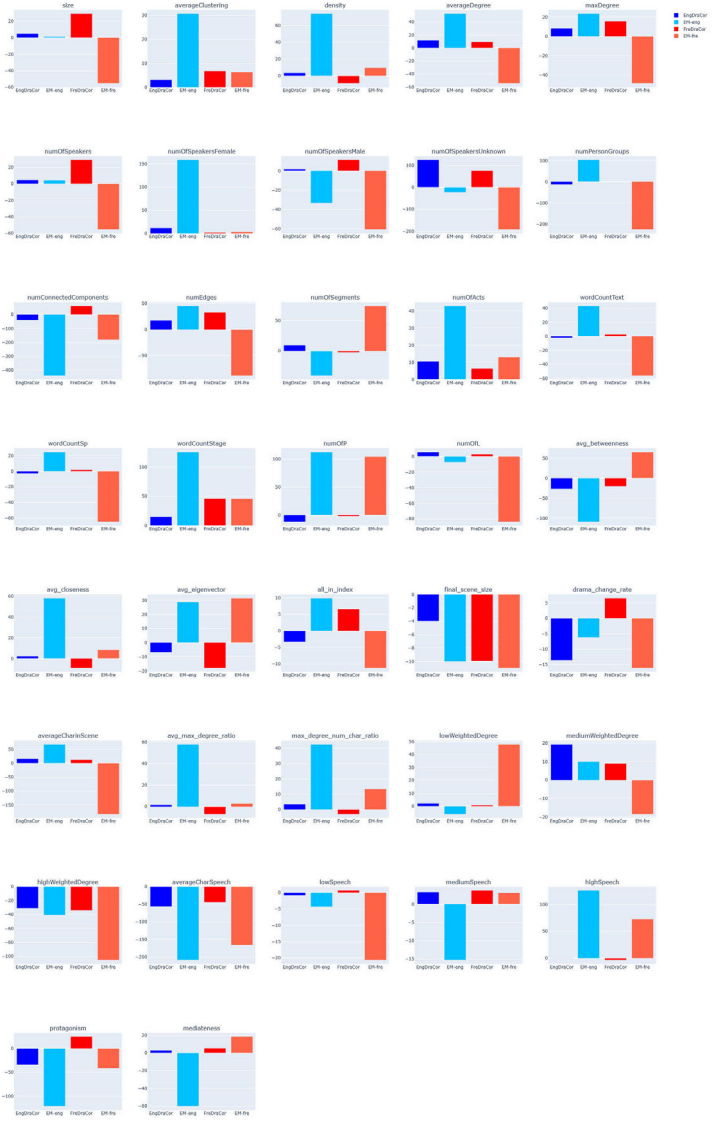


Figure 18: Evolution of the metric 'averageCharSpeech' in the English (1570–1660) and French (1561–1710) DraCor corpora (trendline with 95% confidence interval; extreme outliers removed using the IQR method).

Another phenomenon magnified in both EmDraCor subsets, but also finding confirmation in FreDraCor and EngDraCor, is the decrease in average character speech (Figure 18). The reason for this change becomes clearer if one takes into account the distribution of genres within the corpora. From the first half of the seventeenth century, both of them feature an increasing number of comedies – a form that naturally favours shorter, wittier exchanges over the lengthy tirades typical of high

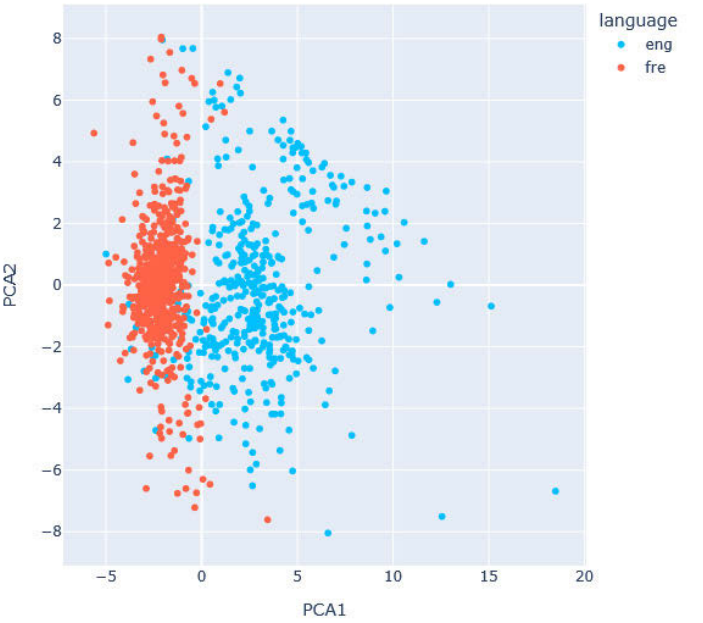
tragedy. As comedies come to represent a larger share of the dataset, it is not unlikely that their characteristically brief speeches progressively pull down the overall average for the metric.

In sum, results from this reproduction experiment return a mixed picture. To begin with, data show that EmDraCor-Eng is actually closer to EngDraCor, as the general trends of the metrics (positive/negative) correspond in 62% of the cases. By contrast, less than 30% of the EmDraCor-Fre metrics exhibit the same trend as in FreDraCor. This discrepancy can be explained by two factors: some known issues in the markup quality of FreDraCor, which influence various network-related metrics and whose impact cannot at this stage be quantified more clearly, and the idiosyncratic composition of EmDraCor-Fre.[14] When trends diverge, the larger corpora nearly always exhibit moderate shifts, indicating that more pronounced increases or decreases in the smaller samples could be reclassified as typical small-sample bias. In only two instances ('size' and 'protagonism' in the two French corpora) are trends significantly opposed; the divergence can be traced to the factors discussed above.

In conclusion, the results from EmDraCor are not fully replicated in the larger corpora, and the picture of dramatic evolution emerging from the corpus cannot be considered statistically robust in absolute terms. The trends it captures are, however, best understood as amplified echoes of those present in the wider dramatic production of the period – exaggerated by the small sample, but not invented by it. In this sense, EmDraCor acts as a sort of 'weather vane', suggesting the main lines of formal development in European theatre even if it cannot prove them beyond doubt. Future improvements in corpus size and markup quality, especially for FreDraCor, will allow these findings to be tested more rigorously.

14 It should be noted that EmDraCor-Fre was not entirely sampled from FreDraCor; one fifth of the texts – mostly early ones – was added through manual encoding.

Figure 19: Principal Component Analysis, English and French DraCor data, 1561–1710.



Ultimately, this reproduction experiment also underlines the viability of vectorisation as a computational method for exploring dramatic forms. As seen in EmDraCor, the compression of plays into meaningful vectors of formal features appears a useful tool for assessing differences between extensive dramatic collections. A contrastive PCA of EngDraCor and FreDraCor (Figure 19) demonstrates this, since it not only displays the clear divide between the two traditions, but also underscores the 'regularity' of French theatre (as testified by a more pronounced clustering) as against the 'irregularity' of the English one (marked by sparse structure with many outliers).