Peer Trilcke, Evgeniya Ustinova, Ingo Börner, Frank Fischer, and Carsten Milling

Detecting Small Worlds in a Corpus of Thousands of Theater Plays

A DraCor Study in Comparative Literary Network Analysis

Abstract: Although homogenized TEI corpora of theater plays from different languages are becoming more and more available, research on plays with a comparative angle is still rare in the field of Computational Literary Studies (CLS). At the same time, approaches of formal network analysis in particular bear huge potential for comparative research due to their modeling of texts as asemantic structures. An attempt to integrate the paradigm of such a formal analysis with general network research on the one hand and literary history on the other hand is the approach of a typification of networks with respect to structural properties such as the the "Small World" concept. However, studies have so far remained limited to smaller and monolingual corpora. In this study, we report on the implementation of different operationalizations of the "Small World" concept and their applications to a corpus of almost 3000 plays. Looking at the results of these analyses, we examine how the different operationalizations of the "Small World" concept relate to each other and discuss how they could be used for a network-based typology of dramatic forms. We finally develop initial ideas for a network-grounded history of dramatic forms in a transnational perspective.

1 Introduction and Research Agenda

1.1 Literary Network Studies

Modeling and analyzing fictional artifacts – such as epics, novels, movies, or plays – as networks has become a widespread procedure in computational humanities (Labatut and Bost 2019). Accordingly, the extraction of network structures from those artifacts and their analysis is not only a challenge for computer scientists (Elson et al. 2010; Lee and Yeung 2012; Agarwal et al. 2013; Krug 2020), but is also being conducted within the scope of decidedly humanities-related research. Thus,

Peer Trilcke, Ingo Börner, Carsten Milling, University of Potsdam
Evgeniya Ustinova, Charles University Prague
Frank Fischer, Free University of Berlin

https://doi.org/10.1515/9783111071824-002

for example, the importance of the Horatio character for the structural stability of the network in Shakespeare's *Hamlet* was described (Moretti 2011, pp. 4–5) or the role of the Mouse character in *Alice's Adventures in Wonderland* as a connector between many characters (Agarwal et al. 2012, pp. 93–94). The networks of Homer's *Illiad* and of his *Odyssey* have been (frequently) investigated (Kydros et al. 2015; Miranda et al. 2018), and there are studies of the amalgamated network of the *Íslendinga sögur* and its subcomponents (Mac Carron and Kenna 2013) as well as of community structures in *Les Misérables* (Newman and Girvan 2004, pp. 12–13). The network-based distinctions between rural and urban novels have been discussed (Elson et al. 2010; Jayannavar et al. 2015), and network measures have been tested as a possible input for genre classification tasks (Coll Ardanuy and Sporleder 2015; Hettinger et al. 2015) and as measures of similarity (Reger 2016). Several papers have also proposed network-based concepts of 'protagonism' as well as quantitative classifications of the characters of fictional texts (Park et al. 2013; Algee-Hewitt 2017; Fischer et al. 2018; Krautter et al. 2018).

So, while network analysis has already proven to be insightful for literary studies, some research has also shown that the "structural intuition" (Freeman 2004, p. 4) of network-based approaches and the related high level of formalization hold particularly high potential for comparative analyses. This is because the abstract modelling and the possibility of a (more or less) 'purely' topological comparison makes the comparative analysis highly independent of any domain- or language-specific origins of the networks.

Three systematic approaches can be distinguished here: first, a comparative approach is possible in the sense of *general network sciences*, in which a universal morphology of all possible forms of networks is elaborated. Second, there are what might be called *cultural* or, third (and more narrowly), *literary network studies*, which focus on the analysis of cultural products, fictional artifacts or even literary works, though against the backdrop of general network science, whose research especially on universal network morphology provides an important theoretical frame of reference. In this sense, networks of literary works and other fictional worlds are regularly discussed with respect to universal topological structures. Examples include studies on mythological networks (Carron and Kenna 2012), on the Marvel universe (Alberich et al. 2002) or on generative approaches to literary network structures (Moretti 2020).

1.2 Studying Small Worlds (in Literature)

One type of network structure that has driven and inspired comparative research is the so-called "small world network" (Watts and Strogatz 1998; Watts 1999a, 1999b), which is defined by a twofold difference: it differs from regular network types as well as from random network types by certain topological characteristics and is suspected "to be widespread in biological, social and man-made systems" (Watts and Strogatz 1998, p. 442). Small world networks seem to be a perhaps universal form that empirical networks tend to take (in contrast to strictly mathematically constructed ones).

It does not come as a surprise that the small world concept has also been used in the computational analysis of literary texts. However, these studies are usually limited to stating that the network constructed for the analyzed work has topological characteristics similar to small world networks – and is therefore structured like a real-world network (Mac Carron and Kenna 2013; Kydros and Anastasiadis 2014; Miranda et al. 2018). In contrast, Stiller et al. (2003), guided by an elaborated anthropological theory, used the small world concept in their analysis of ten Shakespearean plays, in which they not only proved that the plays can be described as small worlds, but also offered evolutionary hypotheses as an explanation for this specific structure. However, even in this approach, which is further developed in Stiller and Hudson (2005), the focus of the analysis (following the paradigm of evolutionary history) lies on the similarity between real-world networks and small world-structured plays.

As interesting as such analyses may be from the perspective of general network science or in light of evolutionary anthropology, from a literary studies point of view they lack at least the reflection of literary form, which has to be considered a major determinant of the networks we extract from literary texts. If, in contrast, one takes into account that the fictional worlds of literature only exist in their mediation through form, there are strong arguments for conceptualizing networks extracted from literary texts first as phenomena of form and not so much as representations of real-world social structures. In this sense, Trilcke (2013, pp. 223–226) has suggested that methods of network science could be applied in literary studies in the context of analyzing form, referring in his own analyses to Klotz's morphological typology of drama (Klotz 1969).

With their corpus-based analysis of more than 450 German-language plays, Trilcke et al. (2016) have taken up this form-oriented approach of literary network analysis and related it to the small world concept. The *small world test* (SWT) proposed by them, a procedure for the algorithmic classification of small world characteristics in the plays of a corpus, aims at analyzing the structural composition of plays and thus could be understood as a network-based contribution to the typology of dramatic forms. For the following re-implementation of this calculation method and the comparison with other implementations, it is crucial that Trilcke et al. (2016) designed their small world test as an outlier test with the aim of identifying particularly exceptional structures in the history of theater plays.

In the present study, we are building on this typology-oriented, form-based approach to small world phenomena in literature. However, we do not want to introduce an additional operationalization of the small world concept for literary studies. Instead, we will, on the one hand, re-implement the outlier-based approach of Trilcke et al. (2016) and, on the other hand, relate it to a much broader operationalization of the small world concept, which Humphries and Gurney (2008) designed to describe the so called "small-world-ness" (SWN) of networks. In this context, we will show that while the two conceptualizations have some overlap, they also open up two different possibilities for typification: while the *SWT small world type* of Trilcke et al. (2016) describes a historic extreme type of dramatic form, the *SWN small world type* of Humphries and Gurney (2008) identifies a systematic standard type of dramatic form.

Overall, with our study we aim to explore different operationalizations of the small world concept as applied to dramatic texts to better understand the conceptual logic behind these different operationalizations. In doing so, we are less concerned with deciding which of the two conceptualizations is the "more correct" one – this would be, in our view, a theoretically fruitless question. Rather, our aim is to contribute to a network-based, structure-oriented literary history of the dramatic form by uncovering and discussing the quite different descriptive potentials of the operationalizations in question.

To be concrete, in the following we will re-implement the above-mentioned small world test (Trilcke et al. 2016) and apply it to a very big drama corpus (VeBiDraCor) with nearly 3,000 structurally fairly homogeneous plays, written in more than ten languages. In addition, we will also implement the scale free test performed by Trilcke et al. (2016) (with reference to Albert and Barabási (2002)), which describes a more strictly defined variant of small world networks. Simultaneously, we will implement the measure of "small-world-ness" (SWN) proposed by Humphries and Gurney (2008), which follows a different conceptual logic than the small world test. Our DraCor-based research corpora (Fischer et al. 2019) as well as both concepts and their corresponding implementations are described in chapter 2. In chapter 3, we will discuss the results of our analysis with regard to the potentials that both small world concepts offer for a comparative typology of dramatic forms. Instead of an outlook, we will eventually sketch some ideas about the history of dramatic form from a network science point of view in chapter 4.

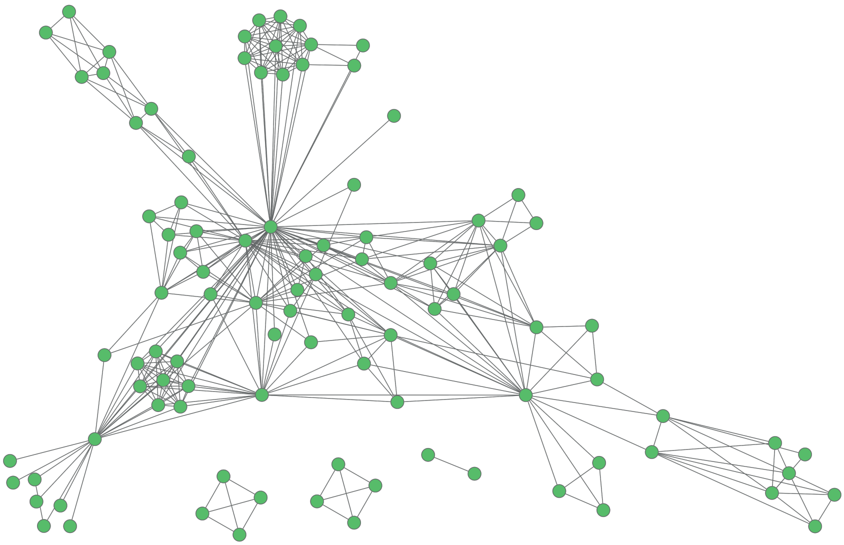

Fig. 1: Johann Wolfgang Goethe, *Götz von Berlichingen mit der eisernen Hand*, 1773.

2 Operationalization and Corpora

2.1 Operationalizing Small Worlds

Small worlds, theorized by Watts as a "deep feature of the social world" that constitutes "a family of graphs" (Watts 1999b, pp. 493–494, 502), are usually characterized in research by two properties, each of which is defined relatively to a "regular" and a "random" network type. Phenomenologically described, small world networks (or more precisely: small world graphs) are, like regular networks, "highly clustered," whereas random graphs are "poorly clustered"; at the same time, small worlds have "small characteristic path lengths, like random graphs" (Watts and Strogatz 1998, p. 440). The description already specifies the two network measures used for the formalization: the network average clustering coefficient C and the average shortest path length L. We will come back to these two measures in a moment.

To begin with a visual representation of the structural properties of small world networks, it is worth taking a look at the network of Goethe's play *Götz von Berlichingen mit der eisernen Hand* (1773) in Figure 1.[1]

1 See https://dracor.org/id/ger000123.

While, on the one side, we see some densely connected areas (“highly clustered”), some very central nodes (such as the character Götz, situated in the upper center) ensure, on the other side, that the paths between these different clusters remain quite short (“small characteristic path lengths”).

This specific structure can be measured. However, before we can apply network measures to dramatic texts, these texts must be *modeled as networks*, where G is a network with n nodes and m edges. This information is extracted from plays encoded in TEI as the following pseudo-structure exemplifies:

```
<div type="segment">
    <sp who="#speaker_A">
        <p>text</p>
    </sp>
    <sp who="#speaker_B">
        <p>text</p>
    </sp>
</div>
<div type="segment">
    <sp who="#speaker_B">
        <p>text</p>
    </sp>
    <sp who="#speaker_C">
        <p>text</p>
    </sp>
</div>
```

While each distinct speaker represents a node n, a relation m is established if the speeches `<sp>` of two or more speakers, assigned via the `@who` attribute (`"#speaker_A"`, `"#speaker_B"`), appear in the same segment `<div>`. The algorithms used for extracting our network data are open-source and can be viewed on our GitHub repositories.[2] It is important to note that this method of extracting networks from dramatic texts is only one possibility among many others and that the scope for interpretation is determined by this mechanism of extraction. At the

2 The xQuery function of the *DraCor API* (https://github.com/dracor-org/dracor-api) *metrics:get-network-metrics* (https://bit.ly/3wDIJ5Y) extracts the segments of a given TEI file using the function *dutil:get-segments* (https://bit.ly/3AWraAK) and for each of these segments gets the distinct speakers with the function *dutil:distinct-speakers* (https://bit.ly/3CIvDYT). The network metrics are calculated based on these extracted features with the *DraCor metrics service* (https://github.com/dracor-org/dracor-metrics) using the Python package *networkx*.

same time, this extraction mechanism has two merits: first, it can be implemented with a manageable effort, and thus can be applied quite quickly to large sets of texts; second, in line with our research interest, it strongly relies on the specific form of the dramatic text. Last but not least, it is of particular importance for our approach that all dramatic networks are constructed in the same way to ensure comparability, which is possible thanks to the homogeneous structure of plays on the DraCor platform (Fischer et al. 2019).

Having extracted dramatic networks from plays in this way, we run a series of small world-related analyses on the data.

First, the small world test SWT proposed by Trilcke et al. (2016): Following the conceptualization by Watts and Strogatz (1998), this test relies on the two above-mentioned measures for network topology, the average clustering coefficient C and the average shortest path length L. We calculate C referring to Watts and Strogatz (1998), while L is implemented as a ratio of the sum of average path lengths for each node and the number of such paths.[3]

For SWT, we assume that each corpus Cor is a set of n plays P, to each of which corresponds a network G. We first calculate C and L for each G. For each G we then generate 1,000 random graphs (following the Erdős-Rényi model (E–R), which, simply put, distributes all edges according to constant probability between the nodes of a graph, see Erdős and Rényi (1959)), and calculate the mean for C and L, resulting in C_{rand} and L_{rand}. Next, for each P we calculate C^{dev} and L^{dev} with

$$C^{dev} = \frac{C_G}{C_{rand}} \tag{1}$$

$$L^{dev} = \frac{L_G}{L_{rand}} \tag{2}$$

As stated above, dramatic small worlds can in general be understood as a type of dramatic form that is shaped by a characteristic combination of *difference* from (i.e. higher clustering than random networks) and *similarity* to (i.e. average path length similar to random networks) another type. Although with C^{dev} and L^{dev} there are measures of similarity and difference, there is not yet a procedure for classifying small worlds. Watts and Strogatz (1998) provide no rule as to when C^{dev} is high enough and when L^{dev} is low enough to classify a network as a small world. Humphries and Gurney (2008, p. 2) therefore propose a “continuously graded notion of small-world-ness” and base the categorical concept of small

3 For networks with more than one component, only the paths between connected nodes are counted.

world derived from this small-world-ness on a postulated threshold value (see remarks below).

The approach to solving this problem suggested by Trilcke et al. (2016) is rooted in the domain focus of their study design. They assume that threshold values for the difference (or similarity) between C_G and C_{rand} (or L_G and L_{rand}), as indicated by C^{dev} (or L^{dev}), must be obtained by comparing them to the whole domain, i.e. in the present case to the particular corpus studied. So, for all *P* in *Cor*, they calculate C^{dev} and L^{dev} and average them. Correspondingly, for SWT we classify G as a dramatic small world if the following two criteria are met:

- **Criterion I**: C^{dev} of a single *P* has to be significantly higher than $\overline{C}^{dev}_{Cor}$ and
- **Criterion II**: L^{dev} of a single *P* must not differ significantly from $\overline{L}^{dev}_{Cor}$.

To decide if the values are significantly higher resp. do not significantly differ, following Trilcke et al. (2016), we calculate a simple deviation test, where we consider anything above or below *mean* $\pm 2\times$ *standard deviation* to be a significant deviation from the norm. For **criterion I**, this means:

$$C^{dev} > \overline{C}^{dev}_{Cor} + 2\delta \tag{3}$$

And for **criterion II** this means:

$$\overline{L}^{dev}_{Cor} - 2\delta \quad < \quad L^{dev} \quad < \quad \overline{L}^{dev}_{Cor} + 2\delta \tag{4}$$

It is, at this point, important to note again the outlier-oriented approach to the small world concept proposed by Trilcke et al. (2016). Assuming a normal distribution of the plays in the corpus (resp. of the values for their structural properties), according to Equation 3 and Equation 4 about 5% of the plays should be typed as SWT type small worlds, whereas 95% should be within the range *mean* $\pm 2\times$ *standard deviation*.

Second, in addition to SWT, following Trilcke et al. (2016), we carry out a scale-free test SFT, in which we check whether dramatic networks can be typified as scale-free networks following the conceptualization by Albert and Barabási (2002). In the interpretation of Trilcke et al. (2016), scale-free networks fulfill criterion I and criterion II *and* are characterized by having a node degree distribution that follows a power law (**criterion III**). Since strict power law distributions are rare (Broido and Clauset 2019), especially in networks as small as the plays studied, in line with we use the following operationalization in line with Trilcke et al. (2016): for the node degree distribution of each P that meets criterion I and criterion II, we calculate the coefficient of determination R^2 for a) a linear, b) a quadratic, c) an exponential, and d) a power law fit. When R^2 of the power law fit is highest, we consider criterion III to be fulfilled and the play has passed the SFT.

Both tests, SWT as well as SFT, follow, as we have already stated, on the one hand a *categorical* logic: the network of a play is either a SWT small world (or a SFT scale-free network) or it is not. On the other hand, the categorical attribution, which both tests provide, is not an absolute one, but always only possible *in relation* to a specific corpus. These two aspects of our operationalization distinguish them fundamentally from the approach of Humphries and Gurney (2008). The measure for small-world-ness S proposed by them operates, in its conceptual basis, with a "continuously graded notion of small-world-ness" (Humphries and Gurney 2008, p. 2), so that a graph, in fact, can be more or less small-world-ish. Only in a second step do the authors introduce a threshold value $S = 1$ with which their operationalization can also be used for categorical attributions. More significant, however, is the difference between our operationalization and the one proposed by Humphries and Gurney (2008) when it comes to the relational aspect of the term. In Humphries and Gurney's conceptualization, small-world-ness S is an absolute measure that can be calculated without referring to a reference group (a corpus in our case). To be more specific, S – with G as any given graph – is calculated as follows (Humphries and Gurney 2008, p. 2):

$$\gamma_G = \frac{C_G}{C_{rand}} \tag{5}$$

and

$$\lambda_G = \frac{L_G}{L_{rand}} \tag{6}$$

so that

$$S = \frac{\gamma_G}{\lambda_G} \tag{7}$$

Thus, while the operationalization of SWT by Trilcke et al. (2016) proposes a *categorial-relative* term, the measure of small-world-ness S by Humphries and Gurney (2008) proposes a *continuous-absolute* term (for which they at the same time indicate options to be applied in a *categorial-absolute* way, resulting in a small world type we will call SWN). It must be noted, in addition, that Humphries and Gurney's threshold $S = 1$ for categorical attribution aims to show small worlds as a *widespread, general phenomenon*; the operationalization by Trilcke et al. (2016), on the other hand, conceptualizes small worlds as a *rare, structurally exceptional phenomenon*. Correspondingly, we expect a high number of dramatic small worlds of the type SWN, but relatively few of the type SWT (and, consequently, even fewer of the type SFT).

Tab. 1: Research corpora overview.

	Number of		Year				Mean Number of	
Corpus Name	**Plays**	**Authors**	**Min**	**Max**	**Mean**	**SD**	**Speakers**	**Segments**
VeBiDraCor	2978	797	−472	2017	1719	332.2	14.3	19.9
VeBiDraCor_Struc	2327	702	−472	2017	1701	367.1	15.3	24.9
VeBiDraCor_Struc_Hist	2246	690	1508	2017	1766	91.7	15.4	25.3

2.2 Research Corpora

To understand how the different operationalizations behave and what typological and historical conclusions can be drawn from their application, we conduct a series of analyses on different corpora of plays (see Table 1). At the heart of our analyses is *VeBiDraCor* – our very big drama corpus, which we created by aggregating all individual corpora currently available through DraCor.[4] For a more detailed breakdown of the corpora aggregated for *VeBiDraCor* and their basic metadata, see Table 2.

From *VeBiDraCor*, we made two derivations, each of which we used to restrict the heterogeneity of dramatic forms somewhat. *VeBiDraCor_Struc* includes only plays whose number of segments is ≥ 5, so that we exclude plays which, due to their shape (i.e. few segments), tend to have rather exceptional networks. With *VeBiDraCor_Struc_Hist* we add a second step of homogenization by restricting ourselves to plays published or first performed after 1500 ("yearNormalized"[5]).

2.3 Analyses

For all research corpora listed above, we

- calculate the small-world-ness measure S;
- perform a categorical application of the small-world-ness measure S, using the threshold = 1 $\rightarrow$ SWN (Humphries and Gurney 2008, p. 2);
- perform the small world test $\rightarrow$ SWT;
- perform the scale-free test $\rightarrow$ SFT.

4 *VeBiDraCor* was created on August 09, 2022 using a dedicated, fully functional Docker image of DraCor (incl. metrics services and API functions), which we also use to version the state of the corpora at a given point in time to provide the identical data and API functionalities for use in replication studies. For more info, see at the end of the chapter.

5 On calculating the "Normalized Year," cf. https://dracor.org/doc/faq.

Tab. 2: Composition of VeBiDraCor.

Corpus			Number of		Year				Mean Number of	
Acronym	**Name**	**Title**	**Plays**	**Authors**	**Min**	**Max**	**Mean**	**SD**	**Speakers**	**Segments**
AlsDraCor	als	Alsatian Drama Corpus	25	13	1816	1937	1914	24.8	12	16.2
BashDraCor	bash	Bashkir Drama Corpus	3	2	1917	1975	1952	30.8	18.7	3.7
CalDraCor	cal	Calderón Drama Corpus	205	1	1623	1760	1663	32.6	18	2.3
FreDraCor	fre	French Drama Corpus	1556	395	1170	2017	1747	85.6	10	19.3
GerDraCor	ger	German Drama Corpus	589	220	1650	1947	1837	57.3	22.3	24.8
GreekDraCor	greek	Greek Drama Corpus	39	4	−472	−388	−425	21.9	11.2	12.7
HunDraCor	hun	Hungarian Drama Corpus	41	25	1558	1970	1860	77.9	24.3	22.9
ItaDraCor	ita	Italian Drama Corpus	139	49	1449	1933	1657	123.2	10.9	26.1
RomDraCor	rom	Roman Drama Corpus	36	3	−215	96	−121	111	11.2	16.3
RusDraCor	rus	Russian Drama Corpus	212	56	1747	1947	1848	51.1	17.5	25.1
ShakeDraCor	shake	Shakespeare Drama Corpus	37	1	1591	1613	1600	6.5	38.7	20.8
SpanDraCor	span	Spanish Drama Corpus	25	9	1878	1945	1913	18.8	22.1	33.8
SweDraCor	swe	Swedish Drama Corpus	68	20	1880	1900	1889	5.7	11.6	14.4
TatDraCor	tat	Tatar Drama Corpus	3	1	1908	1912	1910	2.1	10	12

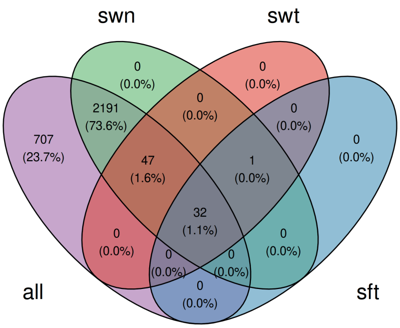

Fig. 2: Relations of the sets of small worlds, with "all" = *VeBiDraCor* and "swn," "swt" and "sft" as the corresponding subsets.

Analyses were carried out based on the output of the DraCor API using a purpose-developed R script.[6]

3 Discussing Results from a Typological Point of View

In the following, we will present and discuss the results of our analyses and further inspect the received data. In our discussions and inspections, we will focus on the following guiding questions: first, how do, in particular, the two small world operationalizations SWN and SWT– the first from general network science, the second from literary network studies – relate to each other in quantitative and phenomenological respects? Second, if we assume that a larger or smaller group of the dramatic networks in our corpora are typified as small worlds: can certain properties be specified for the dramatic networks that do not fall into this group? In other words: can other network-based types of dramatic form be described relative to the group of small worlds?

3.1 Overview of the Results

Table 3 reports both the mean S as well as the absolute number and the share of dramatic networks typified as small world (or scale-free) networks in the cor-

6 The code for the analyses is online, see note at the end of this article.

Tab. 3: Small world attributions by measure and corpus.

			Amount (share) of		
Corpus Name	**N**	**Mean S**	**SWN plays**	**SWT plays**	**SFT plays**
VeBiDraCor	2978	1.68	2,270 (76.2%)	79 (2.7%)	32 (1.1%)
VeBiDraCor_Struc	2327	1.8	2,071 (89.0%)	70 (3.0%)	31 (1.3%)
VeBiDraCor_Struc_Hist	2246	1.8	1,990 (88.6%)	71 (3.3%)	31 (1.4%)

responding corpora according to the different measures. Figure 2 visualizes the relations of the sets of detected small world networks with respect to the different measures.

We first keep three observations:

– A. As expected, SWN typifies a large part of the corpora as small worlds (between 76.2% and 89%), whereas SWT seems to describe only a very specific type of dramatic network (between 2.7% and 3.3%). SFT is even more rare. We will discuss the striking dominance of type SWN small worlds in the subsequent chapter 3.2.
– B. As can be seen in Figure 2 (and as can partly also be mathematically stated a priori), we do have a subset-inclusion order (i.e. a nested set collection), so that *VeBiDraCor* $\supseteq$ *VeBiDraCor_swn* $\supseteq$ *VeBiDraCor_swt* $\supseteq$ *VeBiDraCor_sft*. In other words, the operationalization SWT proposed by Trilcke et al. (2016) turns out to typify not so much a different type, but rather something like an extreme type of a dramatic small world. We will discuss in chapter 3.3 why this extreme type is nevertheless not (necessarily) identical with plays that have the highest S value.
– C. Homogenizing the corpora in structural terms (see *VeBiDraCor_Struc*) leads to a slight increase in S and in the proportion of plays typified as small world resp. scale-free networks. Combining structural and historical homogenization (see *VeBiDraCor_Struc_Hist*) also has only a small effect. We will come back to this observation in chapter 4.

Before discussing the results in more detail from a typological point of view, let us take a brief and exemplary look at how some of the features of dramatic networks influence the small world typification. Table 4 lists the minimum number of speakers (nodes) as well as the minimum number of segments that were sufficient for a play to be typified as SWN, SWT or SFT. Both speakers and segments are important factors in the construction of the networks. As can be seen, SWN can also be assigned to very small, low-segment dramatic networks (see also our discussion of SWN type as a standard type below). In contrast, SWT and even more

Tab. 4: Lowest number of speakers/segments in a play that passed the test.

	Minimal Number of	
Type	**Speakers**	**Segments**
SWN	4 (39 plays, e.g. ger000237)	2 (23 plays, e.g. fre001424)
SWT	17 (greek000027)	9 (rus000091)
SFT	20 (ita000098)	13 (ger000279)

so SFT are sensitive in particular to the size of a network (number of speakers in the play): while the mean of the plays in VeBiDraCor is 14.3 speakers (see Table 1), the SWT play with the lowest size has a speaker count of 17 (for SFT it is even 20). The number of segments, on the other hand, does not need to be above average: while the mean of the plays in VeBiDraCor is 19.9 segments (see Table 1), the SWT play with the lowest size has a segment count of 9 (for SFT it is 13).

These findings are supported by tests of correlation. There is a positive correlation between the value for small-world-ness S and the number of speakers;[7] and there is a slightly lower positive correlation between the value for small-world-ness S and the number of segments.[8] Since both factors are essential for the specific form of a dramatic network, and since small worlds are a type more likely to be found in larger networks, these correlations are unsurprising. Of greater interest are those cases where there is no direct correlation between, say, the number of speakers and the small world status of a dramatic network. More on this later.

3.2 SWN Small Worlds as a Standard Type of Dramatic Form

For our discussion of the typological potentials of the different small world concepts, we return to the three observations A, B, C above – and begin with A, of which two aspects are noteworthy. First, according to SWN, the small-world-ness of dramatic networks is nothing special, but the rule: described from a network analytical point of view, SWN small worlds are a (or even: the) standard type of dramatic form. Second, however, it is important to note that not all dramatic networks can be classified as SWN small worlds in the sense of Humphries and Gurney (2008), which raises the question of what other characteristics these networks have.

7 Pearson's r = .612 ($p < .001$); Spearman's r = .578 ($p < .001$).

8 Pearson's r = .225 ($p < .001$); Spearman's r = .454 ($p < .001$).

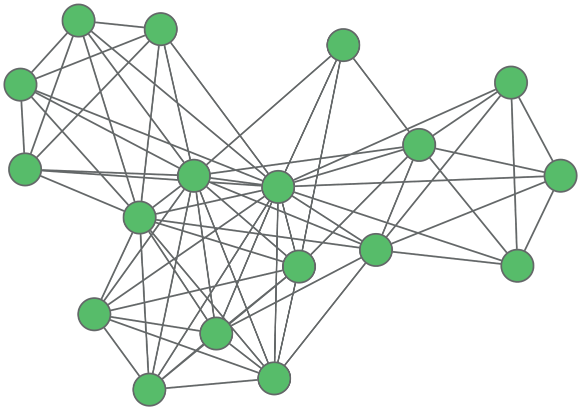

Fig. 3: Georg Kaiser, *Die Koralle*, 1917.

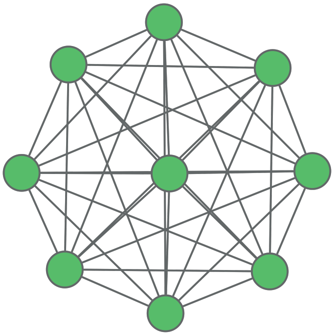

Fig. 4: Calderón, *No hay burlas con el amor*, 1635.

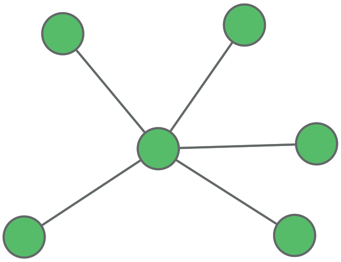

Fig. 5: Louis de Boissy, *L'Apologie du Siècle*, 1734.

To recap: small worlds are a type of network that occupies "a middle ground between regular and random networks," with "high local clustering of elements [...], but also short path lengths between elements" (Humphries and Gurney 2008, p. 1). Let us calculate the average of S for all 2,270 dramatic SWN-small worlds in *VeBiDraCor* and pick the play whose S is closest to this average (which is 1.85), a play called *Die Koralle*, written by Georg Kaiser,[9] which for now shall be taken as our average dramatic SWN small world.

If you look at the respective graph (cf. Figure 3), it actually proves nothing spectacular. *Die Koralle* is a rather average dramatic network, as you get to see it again and again when browsing through DraCor. In this case, we can spot three

9 https://dracor.org/id/ger000545.

somewhat 'more' highly clustered areas (top left, top right, and bottom), and also some nodes in the center of the network that connect these areas, ensuring a rather short path length between them. High clustering, short path length: taken with Humphries and Gurney (2008), this is the core of the small world concept, not more – but not less, either.

Now it might be rather obvious to conclude that the ubiquity of the small world phenomenon, as postulated by general network science, is thus also evident in the dramatic networks that can be typified by SWN. And indeed, there is nothing to argue against this interpretation. From the point of view of literary network studies, however, there is something essential to add. What manifests itself in an average SWN small world network like *Die Koralle* is first of all an effect of *form*, that is: highly clustering areas usually emerge as an effect of the form elements *scene* or *act*. The connection of these areas by central nodes has something to do with the concept of "protagonism," understood as the idea of someone or something carrying the plot that (according to Aristotle) in a play regularly follows the "law of the single strand." In addition, the connections between the areas can also be understood as an effect of the traditional form principle of *liaison des scènes* (even if this principle is not applied in a completely strict sense).

Understanding the structure of dramatic networks (also) as effects of form does not mean to principally reject the views of general network science. It does mean: to speak about dramatic small worlds not as a direct replica of real-world networks, but to always reflect on the literary form as an indispensable mediator of structure, world, society.

And sometimes, form dominates (social) structure. After having discussed small worlds as a form-grounded standard type of dramatic networks, we will now look briefly at the 708 plays from *VeBiDraCor* that were not typified as SWN small worlds (see Table 3). Do the non-small world plays in *VeBiDraCor* share any characteristics, or is every non-small world un-small-world-ish in its own way?

A total of 149 plays from *VeBiDraCor*, for which our analyses did not come to any result at all ("NA") for various reasons,[10] cannot be adequately treated in the context of this question, however. Hence, we set them aside, so 559 plays remain.

Of these, 536 have (at least) two shared characteristics: for their networks, it is both $C = 1$ and $L = 1$, which, third, implies that the density D of the network $= 1$, which in turn indicates the structural property underlying all of this, namely that the graph is fully connected. So, these fully connected (so to say "full den-

10 Almost all reasons have to do with the fact that we have to divide by 0 at some step of the calculation, mostly because the dramatic network is just too small, and as a result L or $C = 0$. Watts also notes that applying the small world concept usually presupposes that a network is "numerically large" (Watts 1999a, p. 495).

sity") graphs are a second type of dramatic form; Figure 4 shows a randomly selected example, *No hay burlas con el amor*, written by Pedro Calderón de la Barca in 1635.[11] Rarely can such dramatic *full density networks* be understood as a kind of reflection of real-world networks; rather, their maximum clustering is primarily an effect of form, typically caused by specific conventions for structuring the scenes in a play (and often determined by the practical conditions of historical types of theater stage).

In the end, 23 plays still remain. Most of them are variants of full density networks, in which only one or very few edges are missing for being fully connected (we call them *high density networks*). Yet there is another small group that can be typified based on a common characteristic. This group of dramatic non-small worlds is the antithesis of the full density networks in that their clustering coefficient is $C = 0$. These are the *dramatic star networks*, as shown in Figure 5 using the example of *L'Apologie du Siècle* by Louis de Boissy, written in 1734.[12] All in all, we have identified eleven dramatic star networks (or slight variants of them) in *VeBiDraCor*.[13]

3.3 SWT Small Worlds as an Extreme Type of Dramatic Form

While the SWN small world type was found to be the quantitatively dominant standard type of dramatic network, the results of the analysis in Table 3 show that the SWT small world type has a very low prevalence. The construction of the SWT measure using an outlier test may suggest that the SWT measure simply typifies those dramatic networks for which the highest small-world-ness value S was calculated in the SWN analysis. This is not the case. In fact, the plays typified by SWT are extreme types of dramatic small worlds, but their identification follows a somewhat different logic than the assignment of very high S values to dramatic networks. In other words, the extreme types typed by SWT are not necessarily the small-worldiest plays in the sense of S. As is shown in Table 5, among the top 30 dramatic networks ranked by S, there are only eight networks (column SWT = TRUE) that are also categorized by SWT as dramatic small worlds.

We, in a first step, choose a phenomenological approach to the different conceptual logics of S and SWT revealed in the ranking and take a look at two

11 https://dracor.org/id/cal000142.

12 https://dracor.org/id/fre000152.

13 These plays deserve to be studied in more detail. Here is a list of their DraCor IDs: fre000029, fre000036, fre000098, fre000152, fre000223, fre000248, fre000317, fre000574, fre001117, fre001511, tat000001.

Tab. 5: *VeBiDraCor* plays ranked by S, top 30.

Rank	Id	Author	Title	Year	S	SWN	SWT	C^{dev}	L^{dev}
1	rus000053	Gogol	Театральный разъезд после представления новой комедии	1842	26.95	TRUE	FALSE	10.5	0.39
2	ger000348	Gleich	Der Eheteufel auf Reisen	1821	14.63	TRUE	FALSE	8.2	0.56
3	fre000709	La Tessonerie	L’Art de Régner	1645	12.58	TRUE	FALSE	6.7	0.53
4	ger000258	Voß	Faust	1823	11.37	TRUE	FALSE	9.6	0.84
5	fre001169	Quinault	La Comédie sans Comédie	1655	11.34	TRUE	FALSE	5.8	0.51
6	fre000518	Dorimond	La Comédie de la Comédie	1662	10.98	TRUE	FALSE	4.9	0.45
7	fre001014	Molière	Le Malade Imaginaire	1673	10.74	TRUE	FALSE	6.4	0.59
8	ger000393	Grabbe	Hannibal	1835	10.64	TRUE	TRUE	10.3	0.97
9	fre000612	Fagan	Momus à Paris	1770	10.29	TRUE	FALSE	4.7	0.45
10	fre001036	Monselet	L’Enfer des Gens de Lettres	1859	10.28	TRUE	FALSE	6.4	0.62
11	ger000123	Goethe	Götz von Berlichingen mit der eisernen Hand	1773	9.79	TRUE	TRUE	9.9	1.01
12	fre001206	Regnard	Le Carnaval de Venise	1699	9.08	TRUE	FALSE	6.5	0.72
13	fre000427	Dancourt	Nouveau Prologue, et Nouveaux Divertissements	1704	8.81	TRUE	FALSE	4.6	0.53
14	ger000302	Schaefer	Faustine, der weibliche Faust	1898	8.52	TRUE	FALSE	6.9	0.81
15	fre000868	Legrand	La Nouveauté	1727	8.2	TRUE	FALSE	5.9	0.71
16	fre000997	Molière	Le Bourgeois Gentilhomme	1671	8.14	TRUE	FALSE	5.9	0.72
17	ger000085	Büchner	Dantons Tod	1835	7.93	TRUE	TRUE	8.8	1.11
18	rus000191	Bulgakov	Война и мир	1932	7.7	TRUE	TRUE	8.2	1.07
19	ger000378	Grabbe	Napoleon oder Die hundert Tage	1831	7.67	TRUE	FALSE	10.7	1.4
20	fre000993	Molière	Les Amants Magnifiques	1670	7.46	TRUE	FALSE	5.4	0.72
21	fre000068	d’Aure	Geneviève ou L’Innocence Reconnue Tragédie Chrétienne	1669	7.45	TRUE	FALSE	3.9	0.52
22	shake000037	Shakespeare	Pericles	1609	7.18	TRUE	TRUE	6.9	0.97
23	swe000047	Strindberg	Lycko-Pers resa	1882	7.17	TRUE	FALSE	5.8	0.81
24	ger000201	Goethe	Faust. Der Tragödie zweiter Teil	1832	6.97	TRUE	FALSE	8.3	1.19
25	ger000278	Avenarius	Faust	1919	6.92	TRUE	TRUE	6.4	0.93
26	ger000564	Büchner	Woyzeck	1837	6.86	TRUE	FALSE	5.4	0.79
27	ger000555	Baggesen	Der vollendete Faust oder Romanien in Jauer	1808	6.86	TRUE	FALSE	8.7	1.27
28	ger000532	Wolfram	Faust	1839	6.83	TRUE	TRUE	6.2	0.91
29	ger000149	Sorge	Der Sieg des Christos	1924	6.76	TRUE	FALSE	5.3	0.79
30	fre001180	Quinault	Roland	1685	6.68	TRUE	TRUE	6.2	0.93

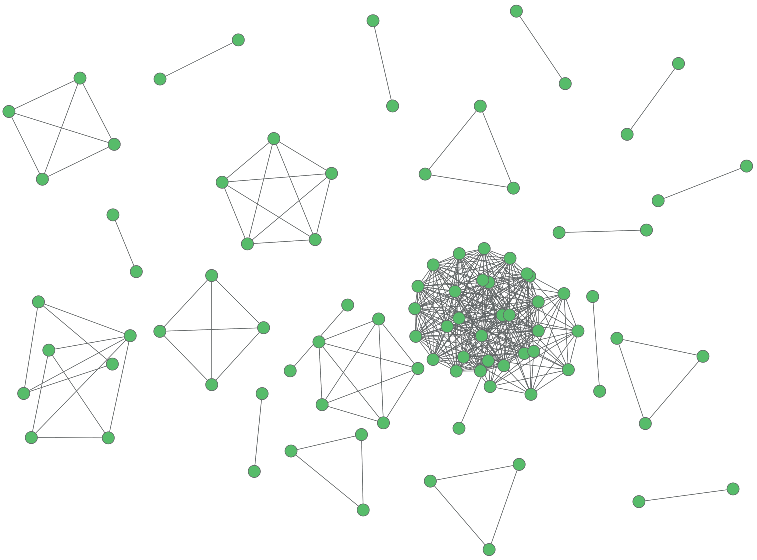

Fig. 6: Nikolaj Vasilevich Gogol, *Театральный разъезд после представления новой комедии*, 1842.

network visualizations. Figure 6 shows the play *Театральный разъезд после представления новой комедии* (*Leaving the Theater after the Presentation of a New Comedy*) by Nikolai Gogol, published in 1842;[14] Figure 7 shows the highest-ranked play also typified according to SWT as a dramatic small world, Christian Dietrich Grabbe's *Hannibal* from 1835.[15] Both dramatic networks are complex, and both dramatic networks are notable at first glance for their numerous semi-autonomous clusters. However, while in Gogol's play the majority of these clusters are unconnected components, in Grabbe's play there are only four separate components in total. Gogol's play thus almost prototypically exemplifies what Watts has called the "caveman graph" (Watts 1999a, p. 500). Yet a caveman network is not a small world.

In contrast, the dramatic network for Grabbe's play (Figure 7) also breaks down into several components. Overall, however, its combination of high clustering and short path length (generated by the bridging edges between the components) makes it much more small-world-ish in the SWT sense.

14 https://dracor.org/id/rus000053.
15 https://dracor.org/id/ger000393.

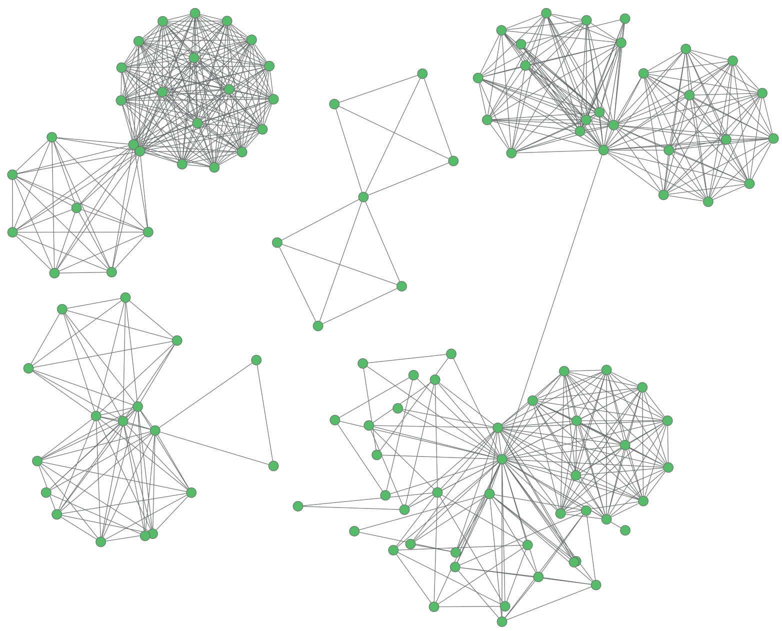

Fig. 7: Christian Dietrich Grabbe, *Hannibal*, 1835.

So, we can note from phenomenological inspection already that our implementation of Humphries and Gurney (2008)'s small-world-ness measure S seems to be quite insensitive to the disintegration of small world candidates into e.g. caveman graphs. But what is the computational reason that Gogol's play has extremely high small-world-ness according to SWN resp. based on S, but is not typified as dramatic small world according to SWT?

The answer lies in the different way the two small world characteristics (high clustering; short average path length) are implemented in the measures. In SWT, which was constructed as a test where both characteristics are checked independently, both characteristics must be met (see equation 3 and equation 4). Humphries and Gurney (2008), on the other hand, relate the two measures to each other by division to create the integrative measure S (see equation 7). However, this makes the small-world-ness measure S susceptible to peaks of one characteristic, which may allow the other characteristic behaving inconspicuously at the same time. To give an example: on the basis of S, a dramatic network may be classified as small world if it has an extremely high clustering (one characteristic with extreme value fulfilled), but at the same time only a medium path length (the other characteristic not fulfilled). This exactly is the case with Gogol's play (as well as with many other top-ranked plays according to S, which SWT does not typify as small world): here, the L^{dev} deviates too much to pass the criterion II test (see equation 4) and thus cannot be typified as a dramatic small world according to SWT.

What follows from this high sensitivity of S to the usefulness of the measure from the point of view of a general network science cannot be discussed here. From a literary network studies perspective, we suggest taking the two measures as a starting point for quite different directions of research. On the one hand, the idea of a measure of small-world-ness may be an interesting starting point for developing a *general morphology of dramatic networks* in which – as outlined above – network structures are described and interpreted as effects of form. In this direction, small-world-ness could turn out to be something like *a general, at least transhistorical and transnational form property of dramatic networks*. On the other hand, dramatic networks of the SWT small world type offer an approach for *a network-based account to dramatic genres*, with genres understood as*historical forms* that, accordingly, emerge under certain historical conditions – and may disappear again.

4 Instead of an Outlook: Some Thoughts on a Network-Grounded History of Dramatic Form

In a last step, we want to at least briefly discuss the potentials of the considerations just presented for research questions on literary history, namely with reference to a set of diagrams showing historical distributions. We first look at how the different network types that we typified in our analysis (i.e. SWN, SWT, SFT) are distributed historically. Afterwards, we further look into the historical distribution, taking into account that the small worlds come from different DraCor corpora (which are merged in *VeBiDraCor*), usually differentiated by national language

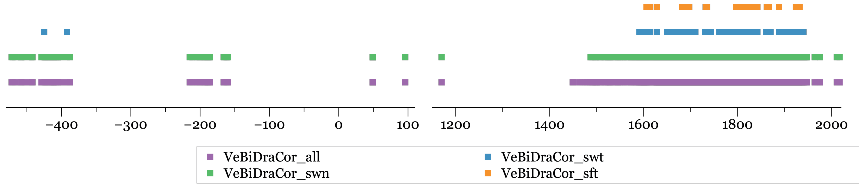

Fig. 8: Historical distribution of dramatic small worlds by different measures.

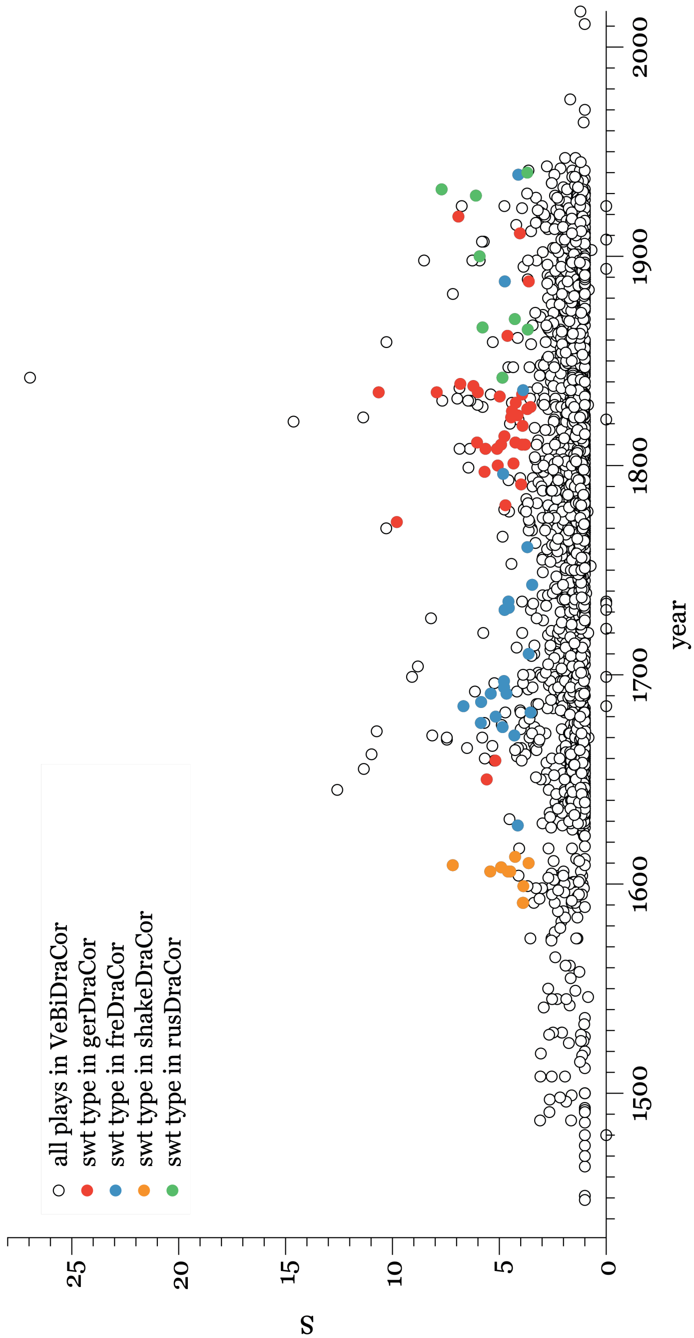

Fig. 9: Historical distribution of all modern SWT type dramatic small worlds differentiated by sub-corpora.

Figure 8 allows a comparison of the historical occurrence of the dramatic networks typified according to the three different concepts. While the purple "baseline" shows all plays (each represented as small square) from VeBiDraCor, the green line represents dramatic small worlds of the SWN type, the blue line represents dramatic small worlds of the SWT type, and the orange line represents dramatic networks of the SFT type. As we have already noted, dramatic small worlds of the SWN type can be understood as a historically more or less indifferent standard type; they describe a transhistorical form option of dramatic networks rather than a specific historical phenomenon.[16] In contrast, dramatic small worlds of the SWT type turn out to be a primarily modern phenomenon. With the exception of two plays by Aristophanes (which would be worth a separate study),[17] the first dramatic small worlds of the SWT type do not appear until the end of the 16th century. Not surprisingly, these first dramatic SWT small worlds of the modern era are plays by Shakespeare, in specific, the two parts of *Henry VI*.[18] Furthermore, the dramatic scale-free networks of the SFT type, additionally characterized by a power-law distribution of the node degree distribution (see criterion III, defined in chapter 2.1), prove to be a genuinely modern phenomenon, at least on the basis of *VeBiDraCor*. Here, too, Shakespeare is the first: the networks of plays like *Timon of Athens*[19] and *Antony and Cleopatra*[20] turn out to be the first examples of dramatic scale-free networks in history. The extraordinary impact of Shakespeare on the history of dramatic form, which is repeatedly attested to him (just to mention studies on the open form in Klotz (1969)), seems to be confirmed by network-grounded analyses.

This, however, is where further research would now have to start, looking much more closely at the structures of these networks, at their realization through dramatic forms and at their relation to dramatic worlds. At this point, we would like to add just one more observation.

Figure 9 shows all dramatic networks from *VeBiDraCor* since 1450, with the small-world-ness value S plotted on the y-axis and the year on the x-axis. The diagram suggests one historical hypothesis: could it be that the form of dramatic net-

16 At this point, at the latest, we however must emphasize that due to the composition of our corpus, we remain limited to a Western perspective in our analyses and thus in our findings. We would be very interested in breaking this Western perspective and including other, non-Western traditions of theatrical texts and so begin to argue from a genuinely global point of view.

17 See https://dracor.org/id/greek000027 and https://dracor.org/id/greek000032.

18 See https://dracor.org/id/shake000020 and https://dracor.org/id/shake000021.

19 See https://dracor.org/id/shake000029.

20 See https://dracor.org/id/shake000035.

works founded by Shakespeare spread in a wave-like fashion through European literatures?

Our analyses, of course, do not provide an adequate answer to this question. In fact, *VeBiDraCor* is far from balanced enough to do so. Moreover, many of the corpora we have brought together in *VeBiDraCor* assemble canonic plays. For a truly comprehensive transnational research on a network-based history of dramatic forms, it turns out, even the already quite large *VeBiDraCor* is still far too small.

Acknowledgment: We warmly thank all the members of and contributors to the DraCor community whose work on the corpora made *VeBiDraCor* – and thus this study – possible in the first place. An overview of all participants can be found at https://dracor.org/doc/credits and, with regard to the TEI files, in the corresponding repositories on GitHub, see https://github.com/dracor-org.

DraCor has, in the context of the project "Computational Literary Studies Infrastructure. CLS INFRA," recently received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101004984.

Data and Code: Data and code of this study are published on GitHub: https://github.com/dracor-org/small-world-paper. The *Very Big Drama Corpus (VeBiDraCor)* was compiled from data published on DraCor.

We conducted the analysis using Docker containers based on an image of *VeBiDraCor* (*vebidracor-api:3.0.0*) (see https://github.com/dracor-org/vebidracor) and *RStudio* as research environment. Instructions on how to setup the Docker containers to reproduce the results using the provided Docker images published on Dockerhub can be found in the README file in the `publication-version` branch.

Calculations were made using a script written in R, which consists of three parts: reading all necessary data and combining it into one dataframe, intermediate calculations of values for the criteria, and checking of conditions of different tests.

References

Agarwal, Apoorv, Augusto Corvalan, Jacob Jensen, and Owen Rambow (2012). “Social Network Analysis of Alice in Wonderland.” In: *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, pp. 88–96. URL: https://aclanthology.org/W12-2513.

Agarwal, Apoorv, Anup Kotalwar, and Owen Rambow (2013). “Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland.” In: *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pp. 1202–1208. URL: https://aclanthology.org/I13-1171.

Alberich, Ricardo, Joe Miro-Julia, and Francesc Rosselló (2002). *Marvel Universe Looks Almost Like a Real Social Network*. DOI: 10.48550/arxiv.cond-mat/0202174.

Albert, Réka and Albert-László Barabási (2002). “Statistical Mechanics of Complex Networks.” In: *Reviews of Modern Physics* 74.1, pp. 47–97. DOI: 10.1103/RevModPhys.74.47.

Algee-Hewitt, Mark (2017). “Distributed Character: Quantitative Models of the English Stage, 1500–1920.” In: *Digital Humanities 2017. Conference Abstracts*, pp. 119–121. URL: https://dh2017.adho.org/abstracts/103/103.pdf.

Broido, Anna D. and Aaron Clauset (2019). “Scale-free Networks Are Rare.” In: *Nature Communications* 10.1. DOI: 10.1038/s41467-019-08746-5.

Carron, Pádraig Mac and Ralph Kenna (2012). “Universal Properties of Mythological Networks.” In: *EPL (Europhysics Letters)* 99.2. DOI: 10.1209/0295-5075/99/28002.

Coll Ardanuy, Mariona and Caroline Sporleder (2015). “Clustering of Novels Represented as Social Networks.” In: *Linguistic Issues in Language Technology* 12. DOI: 10.33011/lilt.v12i.1379.

Elson, David, Nicholas Dames, and Kathleen McKeown (2010). “Extracting Social Networks from Literary Fiction.” In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pp. 138–147. URL: https://aclanthology.org/P10-1015.

Erdős, Paul and Alfréd Rényi (1959). “On Random Graphs I.” In: *Publicationes Mathematicae Debrecen* 6.3–4, pp. 290–297. DOI: 10.5486/PMD.1959.6.3-4.12.

Fischer, Frank, Ingo Börner, Mathias Göbel, Angelika Hechtl, Christopher Kittel, Carsten Milling, and Peer Trilcke (2019). “Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama.” In: *DH2019: »Complexities«. 9–12 July 2019. Book of Abstracts*. DOI: 10.5281/zenodo.4284002.

Fischer, Frank, Peer Trilcke, Christopher Kittel, Carsten Milling, and Daniil Skorinkin (2018). “To Catch a Protagonist: Quantitative Dominance Relations in German-Language Drama (1730–1930).” In: *DH2018: »Puentes/Bridges«. 26–29 June 2018. Book of Abstracts / Libro de resúmenes*. URL: https://dh2018.adho.org/to-catch-a-protagonist-quantitative-dominance-relations-in-german-language-drama-1730-1930/.

Freeman, Linton C. (2004). *The Development of Social Network Analysis: A Study in the Sociology of Science*. Empirical Press.

Hettinger, Lena, Martin Becker, Isabella Reger, Fotis Jannidis, and Andreas Hotho (2015). “Genre Classification on German Novels.” In: *2015 26th International Workshop on Database and Expert Systems Applications (DEXA)*, pp. 249–253. DOI: 10.1109/DEXA.2015.62.

Humphries, Mark D. and Kevin Gurney (2008). “Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence.” In: *PLoS ONE* 3.4. DOI: 10.1371/journal.pone.0002051.

Jayannavar, Prashant, Apoorv Agarwal, Melody Ju, and Owen Rambow (2015). “Validating Literary Theories Using Automatic Social Network Extraction.” In: *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, pp. 32–41. DOI: 10.3115/v1/W15-0704.

Klotz, Volker (1969). *Geschlossene und offene Form im Drama*. 4th edition. Hanser.

Krautter, Benjamin, Janis Pagel, Nils Reiter, and Marcus Willand (2018). *Eponymous Heroes and Protagonists – Character Classification in German-Language Dramas*. URL: https://www.digitalhumanitiescooperation.de/wp-content/uploads/2019/06/p07_krautter_et_al_eng-1.pdf.

Krug, Markus (2020). “Techniques for the Automatic Extraction of Character Networks in German Historic Novels.” PhD thesis. Universität Würzburg. DOI: 10.25972/OPUS-20918.

Kydros, Dimitrios and Anastasios Anastasiadis (2014). “Social Network Analysis in Literature. The Case of ‘The Great Eastern’ by A. Embirikos.” In: *Proceedings of the 5th European Congress of Modern Greek Studies*, pp. 681–702.

Kydros, Dimitrios, Panagiotis Notopoulos, and Georgios Exarchos (2015). “Homer’s Iliad – A Social Network Analytic Approach.” In: *International Journal of Humanities and Arts Computing* 9.1, pp. 115–132. DOI: 10.3366/ijhac.2015.0141.

Labatut, Vincent and Xavier Bost (2019). “Extraction and Analysis of Fictional Character Networks: A Survey.” In: *ACM Computing Surveys* 52.5, 89:1–89:40. DOI: 10.1145/3344548.

Lee, John and Chak Yan Yeung (2012). “Extracting Networks of People and Places from Literary Texts.” In: *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*, pp. 209–218. URL: https://aclanthology.org/Y12-1022.

Mac Carron, Pádraig and Ralph Kenna (2013). “Network Analysis of the Íslendinga Sögur – the Sagas of Icelanders.” In: *The European Physical Journal B* 86.10. DOI: 10.1140/epjb/e2013-40583-3.

Miranda, Pedro Jeferson, Murilo Silva Baptista, and Sandro Ely de Souza Pinto (2018). “The Odyssey’s mythological network.” In: *PLOS ONE* 13.7. DOI: 10.1371/journal.pone.0200703.

Moretti, Franco (2011). “Network Theory, Plot Analysis.” In: *Pamphlets of the Stanford Literary Lab* 2. URL: https://litlab.stanford.edu/LiteraryLabPamphlet2.pdf.

Moretti, Franco (2020). “Simulating Dramatic Networks: Morphology, History, Literary Study.” In: *Journal of World Literature* 6.1, pp. 24–44. DOI: 10.1163/24056480-20201001.

Newman, Mark E. J. and Michelle Girvan (2004). “Finding and Evaluating Community Structure in Networks.” In: *Physical Review E* 69.2. DOI: 10.1103/PhysRevE.69.026113.

Park, Gyeong-Mi, Sung-Hwan Kim, Hye-Ryeon Hwang, and Hwan-Gue Cho (2013). “Complex System Analysis of Social Networks Extracted from Literary Fictions.” In: *International Journal of Machine Learning and Computing*, pp. 107–111. DOI: 10.7763/IJMLC.2013.V3.282.

Reger, Isabella (2016). “Figurennetzwerke als Ähnlichkeitsmaß.” MA thesis. Universität Würzburg. URL: https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-149106.

Stiller, James and Mathew Hudson (2005). “Weak Links and Scene Cliques Within the Small World of Shakespeare.” In: *Journal of Cultural and Evolutionary Psychology* 3.1, pp. 57–73. DOI: 10.1556/jcep.3.2005.1.4.

Stiller, James, Daniel Nettle, and Robin I. M. Dunbar (2003). “The Small World of Shakespeare’s Plays.” In: *Human Nature* 14.4, pp. 397–408. DOI: 10.1007/s12110-003-1013-1.

Trilcke, Peer (2013). “Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft.” In: *Empirie in der Literaturwissenschaft*. Ed. by Philip Ajouri, Katja Mellmann, and Christoph Rauen. Mentis, pp. 201–247.

Trilcke, Peer, Frank Fischer, Mathias Göbel, and Dario Kampkaspar (2016). “Theatre Plays as ’Small Worlds’? Network Data on the History and Typology.” In: *DH2016. Digital Identities: the Past and the Future*. Jagiellonian University, Pedagogical University of Krakow, pp. 385–387. DOI: 10.5281/zenodo.6974706.
Watts, Duncan J. (1999a). “Networks, Dynamics, and the Small-World Phenomenon.” In: *American Journal of Sociology* 105, pp. 493–527. DOI: 10.1086/210318.
Watts, Duncan J. (1999b). *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton Univ. Press.
Watts, Duncan J. and Steven H. Strogatz (1998). “Collective Dynamics of ‘Small-World’ Networks.” In: *Nature* 393.6684, pp. 440–442. DOI: 10.1038/30918.