Network Dynamics, Plot Analysis: Approaching the Progressive Structuration of Literary Texts

Frank Fischer
frafis@gmail.com
National Research University
Higher School of Economics
Russia

Mathias Göbel
goebel@sub.uni-goettingen.de
Göttingen State and University Library, Germany

Dario Kampkaspar
kampkaspar@hab.de
Herzog August Library, Germany

Christopher Kittel
contact@christopherkittel.eu
University of Graz, Austria

Peer Trilcke
trilcke@uni-potsdam.de
University of Potsdam, Germany

Introduction and related works

In recent years, the application of network analysis methods to literary texts has evolved into an independent research field of digital literary studies. Methods for the automated extraction of network data (named entity recognition, co-reference resolution) and their evaluation are of particular importance (Elson et al. 2010; Park et al. 2013; Agrarwal et al. 2013; Fischer et al. 2015; Waumans et al. 2015; Jannidis et al. 2016). Based on the data obtained, several types of analyses were developed: an empirical, quantitative and hierarchical description of literary characters (Jannidis et al. 2016), corpus-based analyses exploring options for historical periodisation of literature (Trilcke et al. 2015) and types of aesthetic modelling of social formations in and by literary texts (Stiller et al. 2003; Stiller & Hudson 2005; Trilcke et al. 2016).

What has been neglected so far (although already suggested by Moretti 2011) is the application of network analysis as a tool for quantitative plot analysis. In fact, current approaches in the field of literary network analysis are not suited to gaining insights into the plot development of literary texts (Prado et al. 2016). The sequential dimension of literary texts, as a consequence of their temporality, usually remains in the dark: what is extracted, visualised and analysed are *static* networks. Plot, however, is essentially a concept supposed to theoretically encompass the temporality of narrative (as well as dramatic) texts: "the repeated attempts to redefine parameters of plot reflect both the centrality and the complexity of the temporal dimension of narrative" (Dannenberg 2005: pp. 435). Plot can be understood as a concept for the description of the "progressive structuration" (Kukkonen 2013: §4) of literary texts.

Research objective: plot analysis

Attempts to further develop literary network analysis towards a quantitative plot analysis must consider the temporal dimension in the modelling of their research objects. The structure of a literary text is to be modelled as a changing sequence of network states. It is only through looking into these network dynamics that we can discuss network-analytic approaches for a quantitative plot analysis.

Following Prado et al. 2016, we are currently extending our research on literary networks (Trilcke 2013; Fischer et al. 2015ff.; Fischer et al. 2015; Trilcke et al. 2016) to the analysis of *progressive* structuration. Our goal is to examine whether (and with what kind of limitations) we can flesh out an operationalisation for the plot analysis of literary texts. By doing so, we are, of course, not trying to replace the semantically rich and versatile concept of 'plot' with a quantitative and thus reductionist concept. Rather, we will show that certain aspects of what is commonly discussed within the framework of plot analysis can be retraced by means of a computer-based analysis of network dynamics.

The visualisation of dynamic graphs, as is common in other domains (Pohl et al. 2008; Frederico et al. 2011), has recently been transferred to literary networks (Xanthos et al. 2016). While it may come in useful when close-reading a text and for didactic purposes, it is unsatisfactory when it comes to an actual corpus-based analysis. There are no canonical methods to help us compare network visualisations generated automatically by force-directed graph drawing algorithms. The reception of dynamic visualisations just does not offer practical analytical means. From a dedicated distant-reading kind of perspective, the calculation of dynamic network metrics and their statistical processing is much more promising as it offers options to describe general characteristics of networks from a larger corpus as well as to compare specific formal types of networks within the corpus.

Measuring dynamic literary networks

A number of basic global measures for the analysis of dynamic networks (i.e., size, density, homogeneity in the distribution of ties, rate of changes in nodes, rate of changes in ties) has been discussed by Carley (2003: pp. 135–36). In addition, Prado et al. 2016 recently suggested the application of actor-oriented measures, especially centrality indices, for the reconstruction of plot development. We are currently applying these measures to [our own corpus](), consisting of 465 German language plays. We also developed a set of other measures with recourse to traditional theoretical concepts for the description of specific phenomena of plot development, especially with regard to types of dramatic expositions (Pfister 1977: pp. 124–36), the "classical" act structure of tragedy and the composition principle of main and secondary plots (Pfister 1977: pp. 286–89). Calculation of these measures was implemented in our own network analysis tool *dramavis* (Kittel/Fischer 2017).

Event-based measures

In general, two types of measures can be distinguished for describing dramatic texts as dynamic networks: event-based and progression-based measures. Event-based measures are used to identify or characterise a particular point in time within a drama. In this context we developed an *all-in index*, a value that identifies the point in time at which all characters have occurred at least once in a drama (see Figure 1).

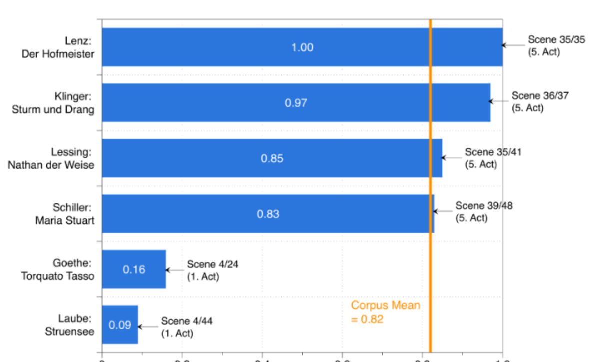

Figure 1: All-in index for 6 selected plays

The *final-scene-size* value characterises a specific point in time, in this case the last scene of a drama. It determines the percentage of all characters of a drama which appear on stage in the last scene. This value shows characteristic differences between dramatic subgenres, especially when comparing tragedies and comedies (see Figure 2).

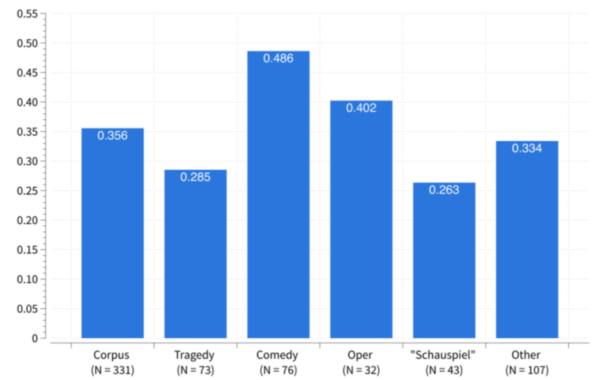

Figure 2: Final-scene-size index (mean values of entire corpus and subgenres)

Progression-based measures

While event-based measures allow assumptions about a particular state/point in time of the dynamic network, progression-based measures allow a general characterisation of the transformation of a dramatic network. In this regard, we introduced a measure we call the *drama-change rate*. The basis of our calculation is a modified Levenshtein distance, which only takes into account insertions ("add character") and deletions ("delete character"). In each case, we compare characters present in two consecutive scenes, eventually resulting in what we call the *segment-change rate* (Figure 3).

Figure 3: Calculation of the segment-change rate (example)

The sum of all *segment-change rates* of a drama divided by the number of all *segment-change rates* is what we call the *drama-change rate*. The development of this rate can be represented in a chart, which due to its shape we tentatively called the *beat chart* (Figure 4).

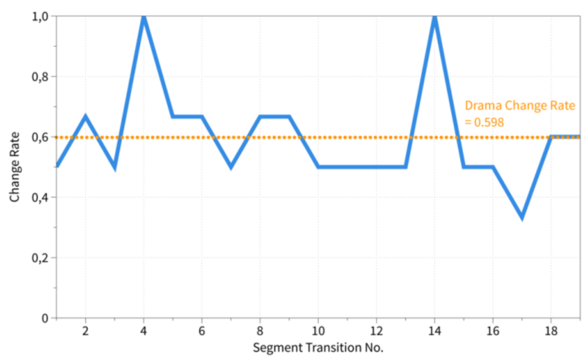

Figure 4: Beat chart for Goethe's play "Iphigenie auf Tauris" (1787)

Having calculated the *drama-change-rate* value for the entire corpus, we can start to compare a larger set of dramas with each other (Figure 5). It becomes evident that our corpus does not exhibit a clear trend along the timeline. Instead, we witness the emergence of characteristic types of dramas, which differ characteristically from the other texts in our corpus.

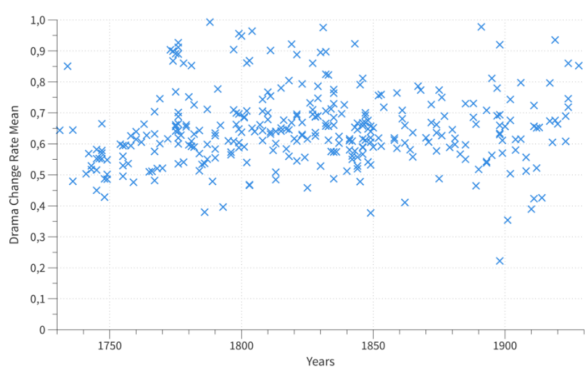

Figure 5: Historical distribution of drama-change rates in the DLINA corpus

On the one hand, we can identify dramas exhibiting a high *drama-change rate*, i.e., highly dynamic dramas with a constant alternation of characters on stage (Figure 6). On the other hand, low-dynamic dramas can be identified, with only small changes taking place between scenes (Figure 7).

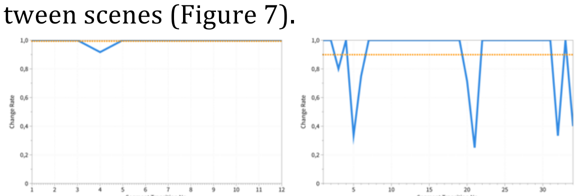

Figure 6: Two examples of highly dynamic dramas – left: Goethe's "Egmont" (1788); right: Lenz's "Der Hofmeister" (1774)

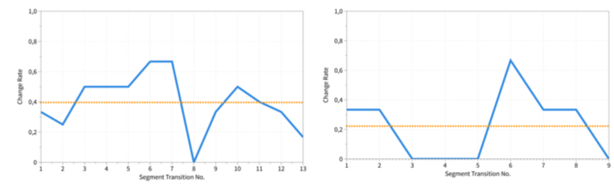

Figure 7: Two Examples for low-dynamic dramas – left: Goethe's "Der Bürgergeneral" (1793), right: Rilke's "Ohne Gegenwart" (1898)

A further option for the comparative analysis of our oscillatory *beat charts* is the analysis of the standard deviation of all *segment-change rates* of a drama. We can again distinguish two particularly striking types: On the one hand, there are dramas with a high standard deviation, indicating that extensive changes of characters alternate with small changes; we can call them high-dynamic dramas (Figure 8). On the other hand, there are dramas showing a low standard deviation, so the change on stage takes place in the same rhythm, we call these texts low-dynamic dramas (Figure 9).

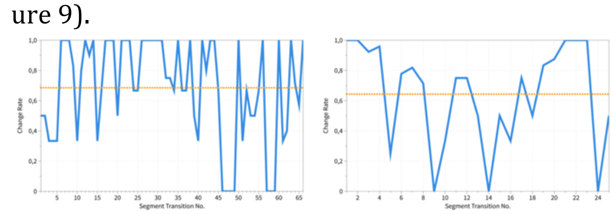

Figure 8: Two Examples for high-dynamic dramas – left: Benkowitz's "Die Jubelfeier der Hölle" (1801), right: Goethe's "Faust I" (1808)

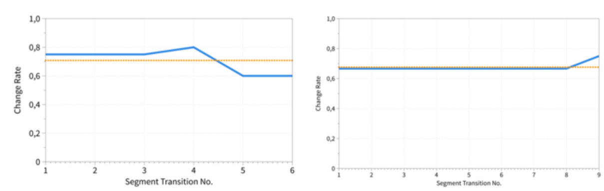

Figure 9: Two Examples for low-dynamic dramas – left: Schnitzler's "Anatol" (1893), right: Schnitzler's "Der Reigen" (1902)

The preceding cases each describe characteristic deviations regarding the *drama-change rate* of groups of texts. In addition, a 'normal type' of drama can be identified, which corresponds to the mean value for the corpus both in terms of arithmetic mean and standard deviation (Figure 10).

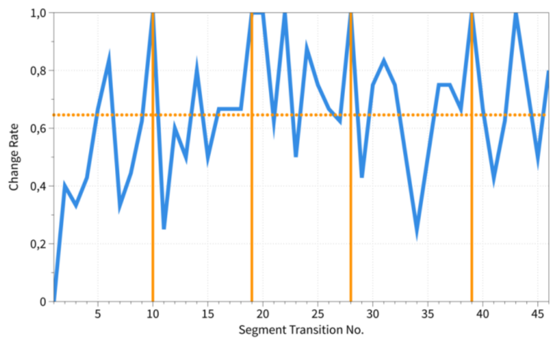

Figure 10: Beat chart for Ganghofer's "Der Herrgottschnitzer von Ammergau" (1880)

It also appears that strong upward shifts, accompanied by a complete exchange of characters on stage, often coincide with act boundaries (vertical orange lines), which is in accordance with historical conventions of dramatic composition.

Summary and further research

Our research on dynamic networks provides basic components for a quantitative analysis of the progressive structuration of dramatic texts. Future research will have to develop and evaluate additional measures and it will be decisive to hold interpretations of these measures against the backdrop of historical drama poetics.

Bibliography

Agarwal, A., Corvalan, A., Jensen, J., Rambow, O. (2012). Social Network Analysis of Alice in Wonderland. *Proceedings of the Workshop on Computational Linguistics for Literature.* Montréal, pp. 88–96, http://www.aclweb.org/anthology/W12-2513 [accessed 27 Mar 2017].

Carley, K. M. (2003). Dynamic Network Analysis. Breiger, R., Carley, K. M., Pattison, P. (eds.): *Dynamic Social Network Modeling and Analysis.* Workshop Summary and Papers. Washington D. C., pp. 133–45, http://www.nap.edu/read/10735/chapter/9 [accessed 27 Mar 2017].

Dannenberg, H. (2005). Plot. Herman, D., Jahn, M., Ryan, M.-L. (eds.): *The Routledge Encyclopedia of Narrative Theory.* London: Routledge, pp. 435–39.

Elson, D. K., Dames, N., McKeown, K. R. (2010). Extracting Social Networks from Literary Fiction. *Proceedings of ACL 2010.* Uppsala, pp. 138–47, http://dl.acm.org/citation.cfm?id=1858696 [accessed 27 Mar 2017].

Federico, P., Aigner, W., Miksch, S., Windhager, F., Zenk, L. (2011). A Visual Analytics Approach to Dynamic Social Networks. *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW).* Graz, http://publik.tuwien.ac.at/files/PubDat_198995.pdf [accessed 27 Mar 2017].

Fischer, F., Göbel, M., Kampkaspar, D., Trilcke, P. (2015). Digital Network Analysis of Dramatic Texts. *Digital Humanities 2015. Conference Abstracts.* University of Western Sydney, http://dh2015.org/abstracts/xml/FISCHER_Frank_Digital_Network_Analysis_of_Dramati/FISCHER_Frank_Digital_Network_Analysis_of_Dramatic_Text.html [accessed 27 Mar 2017].

Fischer, F., Göbel, M., Kampkaspar, D., Kittel, C., Trilcke, P. (2015ff.). [Blog] dlina – Digitally-Driven Literary Network Analysis (of Dramatic Texts). https://dlina.github.io/ [accessed 27 Mar 2017].

Jannidis, F., Reger, I., Krug, M., Weimer, L., Macharowsky, L., Puppe, F. (2016). Comparison of Methods for the Identification of Main Characters in German Novels. *Digital Humanities 2016. Conference Abstracts.* Jagiellonian University & Pedagogical University, Kraków, pp. 578–82 http://dh2016.adho.org/abstracts/297 [accessed 27 Mar 2017].

Kittel, C., Fischer, F. (2017). dramavis (v0.3). GitHub repo: https://github.com/lehkost/dramavis [accessed 27 Mar 2017].

Kukkonen, K. (2013). Plot. Hühn, P. et al. (eds.): *The Living Handbook of Narratology.* Hamburg. http://www.lhn.uni-hamburg.de/article/plot [accessed 27 Mar 2017].

Moretti, F. (2011). Network Theory, Plot Analysis. *Stanford Literary Lab Pamphlets.* No. 2, 1 May 2011, http://litlab.stanford.edu/LiteraryLabPamphlet2.pdf [accessed 27 Mar 2017].

Park, G.-M., Kim, S.-H., Cho, H.-G. (2013). Structural Analysis on Social Network Constructed from Characters in Literature Texts. *Journal of Computers* 8.9, pp. 2442–47, http://ojs.academypublisher.com/index.php/jcp/article/view/jcp080924422447/7672 [accessed 27 Mar 2017].

Pfister, M. (1977). Das Drama. Theorie und Analyse. München: Fink.

Pohl, M., Reitz, F., Birke, P. (2008). As Time Goes by. Integrated Visualization and Analysis of Dynamic Networks. *AVI 2008 – Proceedings of the Working Conference on Advanced Visual Interfaces.* Neapel, pp. 372–75, http://doi.acm.org/10.1145/1385569.1385636 [accessed 27 Mar 2017].

Prado, S. D., Dahmen, S. R., Bazzan, A. L. C., Mac Carron, P., Kenna, R. (2016). Temporal Network Analysis of Literary Texts. *Advances in Complex Systems (ACS)* **19**(3), pp. 1–19, https://arxiv.org/pdf/1602.07275 [accessed 27 Mar 2017].

Rochat, Yannick (2014). Character Networks and Centrality. Thèse de Doctorat, Lausanne, https://infoscience.epfl.ch/record/203889/files/yrochat_thesis_infoscience.pdf [accessed 27 Mar 2017].

Stiller, J., Nettle, D., Dunbar, R. I. M. (2003). The Small World of Shakespeare's Plays. *Human Nature* 14, pp. 397–408, https://www.staff.ncl.ac.uk/daniel.nettle/shakespeare.pdf [accessed 27 Mar 2017].

Stiller, J., Hudson, M. (2005). Weak Links and Scene Cliques Within the Small World of Shakespeare. *Journal of Cultural and Evolutionary Psychology* 3, pp. 57–73.

Trilcke, P., Fischer, F., Göbel, M., Kampkaspar, D. (2015). 200 Years of Literary Network Data [blogpost], https://dlina.github.io/200-Years-of-Literary-Network-Data/ [accessed 27 Mar 2017].

Trilcke, P., Fischer, F., Göbel, M., Kampkaspar, D., Kittel, C. (2016). Theatre Plays as 'Small Worlds'? Network Data on the History and Typology of German Drama, 1730–1930. *Digital Humanities 2016. Conference Abstracts.* Jagiellonian University & Pedagogical University, Kraków, pp. 385–87 http://dh2016.adho.org/abstracts/407 [accessed 27 Mar 2017].

Trilcke, P. (2013). Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft. Ajouri, P. et al. (eds.): *Empirie in der Literaturwissenschaft.* Münster: mentis, 201–247.

Waumans, M. C., Nicodème, T., Bersini, H. (2015). Topology Analysis of Social Networks Extracted from Literature. *Plos One*, 3 June 2015, http://dx.doi.org/10.1371/journal.pone.0126470 [accessed 27 Mar 2017].

Xanthos, A., Pante, I., Rochat, Y., Grandjean, M. (2016). Visualising the Dynamics of Character Networks. *Digital Humanities 2016. Conference Abstracts.* Jagiellonian University & Pedagogical University, Kraków, pp. 417–19 http://dh2016.adho.org/abstracts/407 [accessed 27 Mar 2017].